Modelling of snow depth and snow density based on capacitive measurements using machine learning methods.

Nils-Olav Skeie*, Henrik Nikolai Vahl, Håkon Viumdal

Department of Electrical Engineering, Information Technology and Cybernetics (EIK)
University of South-Eastern Norway (USN), N-3918 Porsgrunn, Norway
*(nils-olav.skeie@usn.no)

Abstract

In countries with cold winters, snowpack will affect the hydropower production during the melting periods. To optimize the hydropower production, it is relevant to consider information from the snowpack to estimate the water content when melting. Several techniques and devices can be used to measure the water content of the snowpack. This paper discusses a prototype based on capacitive measurements with a small footprint, and the development of data driven models to estimate the snow density, snow depth and snow water equivalent in a snowpack. The device was deployed in a snowy area throughout the winter with logging while manual reference measurements were made sporadically. Machine learning methods were used for developing the models, and several models were combined to estimate the water content of the snowpack. The developed model estimated the snow density, snow depth and snow water equivalent during the wintertime with good results. However, during the springtime, the capacitive measurements have some limitations.

Keywords: snow density, snow water equivalent, capacitive sensor, model development, machine learning.

1 Introduction

1.1 Background

Measurement of snowpack has been a challenge for a long time, one of the first research papers handling this challenge was published in the early 80s (Denoth et al., 1984). Snowpack is defined as the mass of snow on the ground that is compressed and hardened by its own weight.

Weather forecast is considered the most important input to the models used for predicting hydropower production. However, in areas where snow accumulates during winter and melts in the spring, the inflow and hence the predictions based on these models, will be highly affected.

Today measurements of snow depth and density are often performed manually at the end of the winter season with the aim of estimating the melting inflow. These measurements must often be taken in remote and impassable locations, making it time-consuming and expensive. As an alternative way of measuring the snow depth and density, an autonomous system with a minimal environmental footprint that can be deployed remotely and transmit data is proposed.

In this work a prototype system was developed, based on capacitance measurements, and deployed in a snowy area in autumn 2020. The prototype was used to record measurement values during the winter period 2020/2021 and these measurements were used as input data for the work of a master thesis at University of South-Eastern Norway (USN) spring 2021. The focus of this master thesis is a modelling approach using machine learning methods (Vahl, 2021).

The focus in this work is for hydropower systems. However, the prototype system is a more general system that can also be applied for other purposes based on snowpacks like measuring skiing conditions and evaluating the risk for avalanches.

1.2 Previous work

Since measuring the snowpack has been a challenge for decades, several works and projects have been performed within this area.

An overview of instruments for measurement of the snowpack is described in (Denoth et al, 1984). The focus of these measurements is to estimate the snow water equivalent (SWE) in the snowpack where SWE is defined as

$$SWE = h_s \frac{\rho}{\rho_w}$$

where h_s is the snow depth, ρ is the density of snow and ρ_w is the density of water, measured in g/cm^3. Measurement of dielectric properties of snow is described in (Hallikainen et al. 1982), and an overview of snow dielectric devices and applications is described in (Denoth and Wilhelmy, 1989).

Newer work about liquid water content in snow is described in (Niang et al., 2006) and (Techel and Pielmeier, 2010).

Some ongoing projects are “Long-term snow water equivalent measurements” at WSL (Wsl, 2021) and “Current snow cover” at CCIN (Ccin, 2021).

The liquid water content in snow is important for hydropower systems as the inflow of water at springtime will be a combination of water from raining and melting of snow. Predicting the amount of inflow is done using models based on among others weather forecasts. However, these models are missing regular inputs from areas where snow accumulates during winter and melts in the springtime.

In Norway, the Norwegian Water Resources and Energy Directorate (NVE) published a report where snow pillows are recommended as automatic snow water equivalent sensors (Stranden et al., 2015). Snow pillows are sensors having a large environmental footprint with a weight and distance sensor to measure the weight and height of the snowpack.

A project was started at USN in 2019, in cooperation with Skagerak Energy AS, to try to develop an autonomous measurement system for snow depth, snow density and snow water equivalent for remote locations, with a small footprint. The first part of the project looked at the system structure of an autonomous measurement system and several possible measurement principles, and a capacitive solution was proposed (Bjerke et al., 2019). A prototype measurement system, based only on capacitance measurements was developed and used for logging (Murillo Abril, 2020), (Murillo Abril et al., 2020). The Covid-19 situation in springtime 2020 with limitation of traveling and the absence of snow at the USN campus limits the number of valid measurements. A new prototype was developed autumn 2020 with five capacitive measurements at different heights in addition to measurement of the atmospheric pressure and temperature. This prototype was deployed in a snowy area, close to Lillehammer (in Norway), throughout the winter season in 2020/2021 with automatic logging of sensor values while manual reference measurements were made sporadically.

1.3 Outline of paper

Section 2 provides a discussion of the system, the prototype developed and deployed, the logging of the sensor values from the prototype, and any manual measurements. Section 3 gives an overview of the machine learning methods. Section 4 gives an overview of the model fitting and validation. The results are discussed in Section 5 and some conclusions are drawn in Section 6.

2 System description

The prototype housing consists of a two-meter heigh grey plastic pipe, with a 90-degree bend on the top. The capacitance sensor devices are located at fixed heights of the plastic pipe, and the temperature and pressure sensor are located at the top of the plastic pipe. The hypothesis for the project was that the snow depth and the snow density could be estimated based on the capacitance measurements at different layers in the snowpack.

A picture of the prototype is shown in Figure 1 covered by about 70 cm of snow, with the two upper capacitance sensors as the black objects pointing to the left from the pipe. The three remaining capacitance sensors are covered by the snow.

Figure 1: The prototype, covered by about 70 cm with snow. The two upper capacitance sensors can be seen as the two black objects pointing to the left of the pipe.

The measurement system of the prototype consists of an Arduino Nano located at the top of the plastic pipe. The Arduino Nano system starts running once a minute, read the sensor values, convert the capacitance sensor outputs to voltage range [0,5] Volt, the temperature to °C and the pressure to mBar, and transmit the converted values over the serial line (USB) on a Modbus based protocol. The capacitive sensor is a “Capacitive Soil Moisture Sensor v1.2” device with a 0-5V interface, the temperature sensor is a TMP36 silicon device with a 0-5V interface, and the pressure sensor is a 4Tech absolute pressure device with a 4-20 mA interface. The 4-20mA interface is converted to a 0-5V interface using a 250 Ohm resistor. However, the pressure sensor must be powered by a separate 12 V DC power supply while the other sensor devices are powered by the 5 V DC output of the Arduino system. The vertical locations of the capacitive sensor devices are 10, 30, 50, 80 and 110 [cm].

A use case diagram, made by Unified Modeling Language (UML), for the Arduino software is shown in Figure 2. As shown in the figure the Arduino software will collect the sensor values and transmit the converted sensor values on the serial line. A remote logging system is needed to record, filter, and store these values.

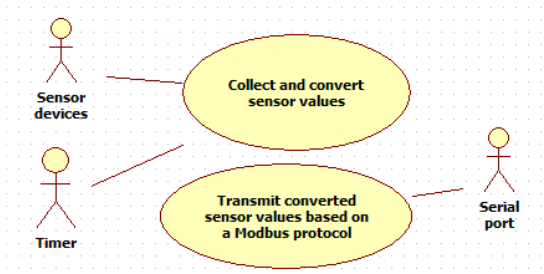

Figure 2: A use case diagram showing the functionality of the Arduino software. Every minute, read the sensor values, convert to the right unit, and transmit on the serial port using a Modbus based protocol.

Machine learning methods were applied, and supervised learning was the selected method as prediction models for the snow depth, snow density and snow water equivalent were needed. Hence corresponding values for snow depth and snow density should be recorded together with the sensor values.

A use case diagram, in UML, for the data storage software (DSS) is shown in Figure 3.

The software has four main functionalities; 1) collect, low pass filter, store and display the values from the measurement system. 2) handle configuration of sensor types, size of low pass filter and how often to store the measurement values on the Comma Separated Values (csv) file. 3) Allow for input of reference values for snow depth and snow density that can be used for the training the models. These values will be stored on the csv file together with the sensor values. 4) Logging of the values on the csv file at specific times independent of receiving data from the measurement system. The default setup was a moving average low pass filter size of 8 and storing the values in csv file every 30 minutes.

The serial port, used for communication between the Arduino system and DSS is an USB port, which also contains the power (5VDC) for the Arduino system.

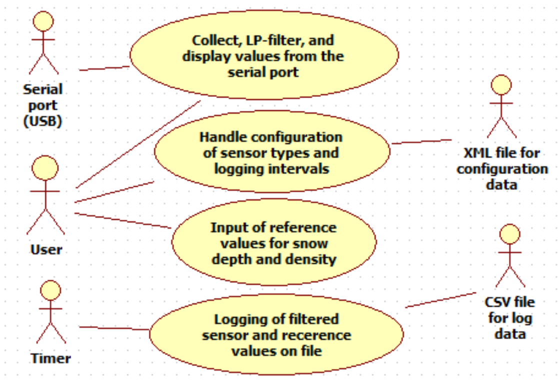

Figure 3: A use case diagram showing the functionality of the logging software, for logging the values from the measurement system.

Figure 4 shows the measuring node to the left, based on a vertical plastic tube with the capacitance sensor devices at fixed height, seen as the white areas in the figure. The lowest capacitance sensor is not visible in this figure. The right side shows the connection between the Arduino system and the DSS, and the protocol used between these systems.

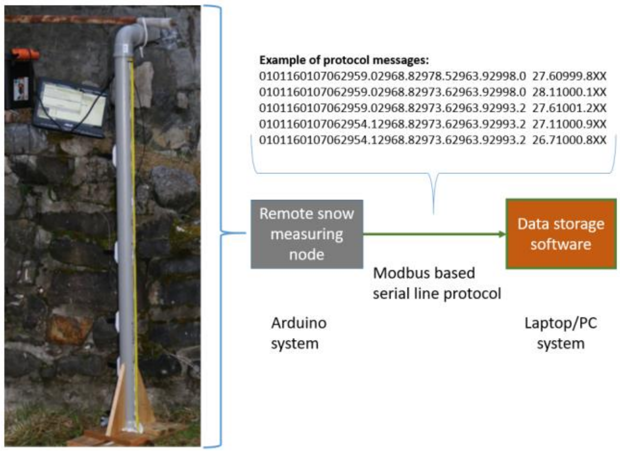

Figure 4: The measurement node to the left with the capacitance sensors at the white areas, and the connection with some examples of the protocol to the right.

The DSS will low pass filter and store the sensor values on a csv file at fixed time intervals, configurable in the DSS. Figure 5 shows a plotting option in DSS when testing manual covering capacitance sensor #2 and #3 with snow. Sensor #2 at 13:40 and sensor #3 at 14:50, with sensor values in mV. Sensor #1 is the lowest sensor device already covered by snow. The change in the voltage is about 100 mV depending on the water content of the snow.

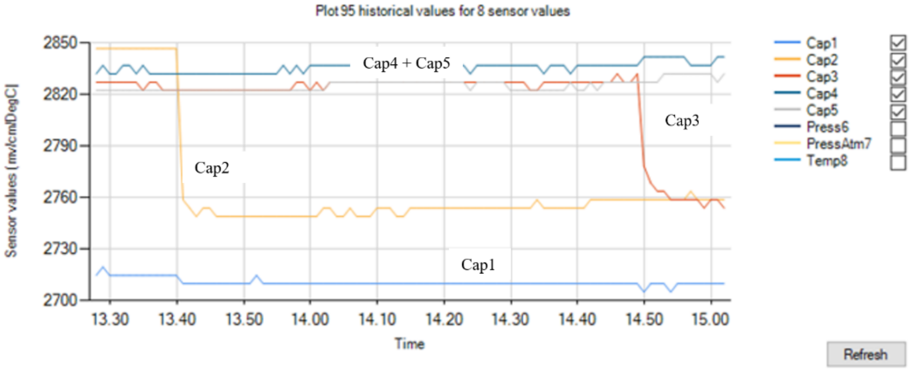

Figure 5: Plotting of the capacitance sensor values when covered by snow. Cap1 is the lowest capacitance sensor, covered by snow. Cap2 is manually covered by snow at 13:40, and Cap3 at 14:50. Cap 4 and Cap 5 is not covered by snow.

The format of the csv file is one line for each measurement containing the time stamp, the values from the capacitance sensors (the lowest one first), temperature sensor, pressure sensors, and any manually measured values for snow depth and snow density. The size of csv file was limited by making a new file for each month.

The manual measurements were made by using a metal pipe with an inner diameter of 6.6 cm inserted into the snowpack, measuring the depth and the weight of the sample. The snow density was calculated based on the pipe diameter and the weight of the snow. All the data needed for the model development are stored in the csv files.

3 Model development

Several tools and frameworks were used for analysis, preprocessing and development of the models. The reason for using several frameworks was to get some experience with the analyzing tools and the process of making models for different frameworks. MATLAB was used for both analyzing, preprocessing and development of the models. ML.NET is a free machine learning framework from Microsoft and was used together with C# for developing models. Keras and TensorFlow was used together with Python to develop models, the tools used in Python was scikit learn. Open Neural Network Exchange (ONNX), an open standard for machine learning models, was evaluated for transferring some of the models between these frameworks but was not used since separate models were developed in each framework.

Supervised learning is used so only the measurement with references can be used for developing the model. The periods for measurements with manual references are 1) 22 to 27-NOV-20, 2) 28 to 31-DEC-20, 3) 12 to 13-FEB-21, and 4) 30-MAR to 2-APR-21. These data were analyzed and some of the data had to be removed because of a power loss error. There is also a challenge

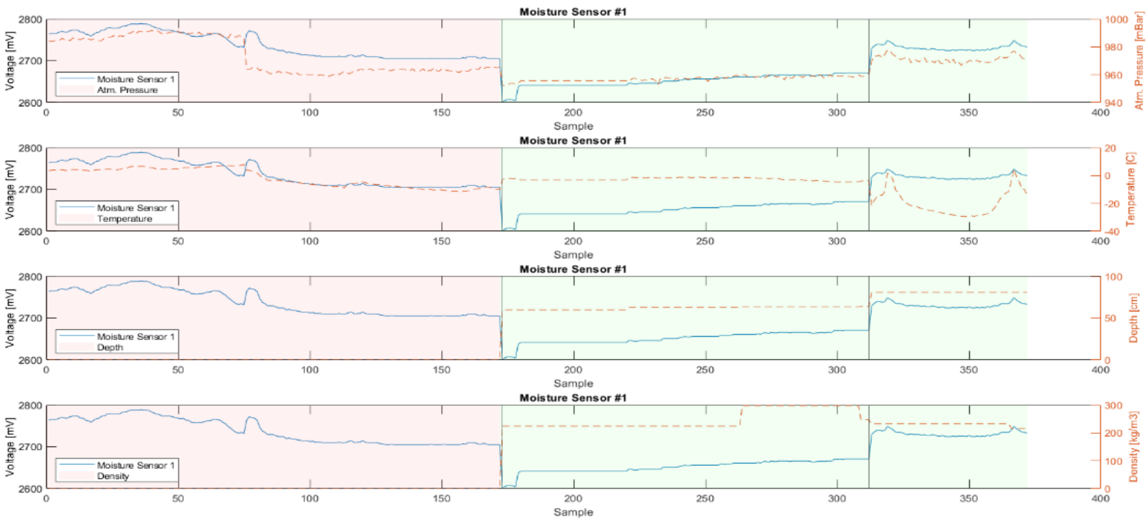

Figure 6: All valid samples for the lowest capacitance sensor with relation to pressure, temperature, depth, and density. Three sections, first section is for November (samples 0 to 172) (without snow), second section for December (samples 172 to 312) and third section for February (samples 313 to 372).

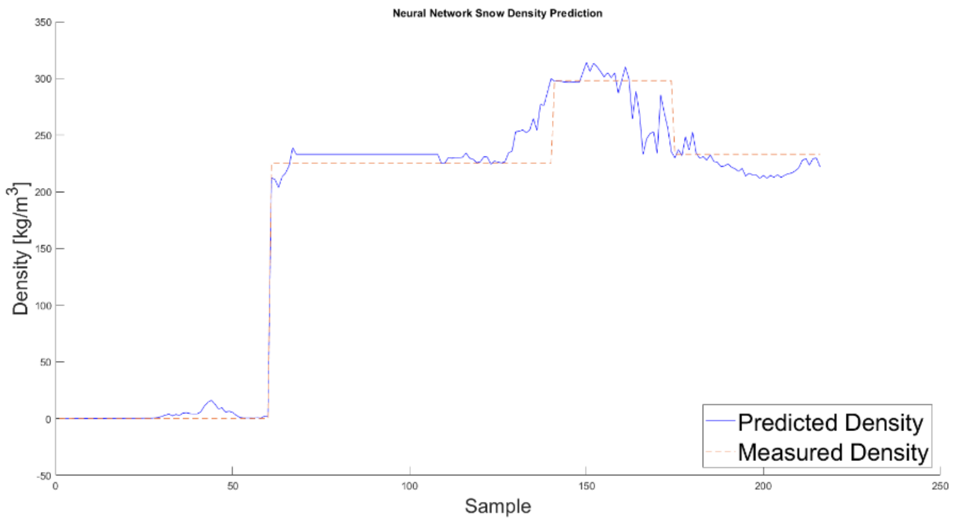

Figure 7: The model for estimating the snow density based on the BR algorithm in MATLAB.

with measurements using a sensor device at fixed location so all the measurements from the last period was removed. This will be discussed as part of the discussion section. The total number of samples that can be used for training of the models are 372.

The valid samples for capacitance sensor #1, the lowest sensor, is shown in Figure 6. The first section, the red section, is the samples for November [1-172], the second section is December [173-312], and the last section is February [313-372]. In each row is the capacitance sensor compared with the pressure, temperature, depth, and density.

A Principal Component Analysis (PCA) was performed indicating that only four principal components are needed to explain 96% of variance in the data. The system has seven variables, 5 capacitance sensor devices, one temperature sensor device and an absolute pressure sensor device. The PCA indicated that all sensor devices were important for the needed information.

The sample set was divided into a training set of 216 samples and a test set of 156 samples.

4 Results

The goal of this project is to develop a model to estimate the Snow Water Equivalent (SWE) parameter for the snowpack, based on the snow depth and snow density. SWE is based on the snow depth and snow density so two separate models must be developed first, one to estimate the snow depth and one model to estimate the snow density.

4.1 Model trained using MATLAB

The Regression Learner app was used in MATLAB to train the models with different algorithms. The best model suggested by the Regression Learner app was

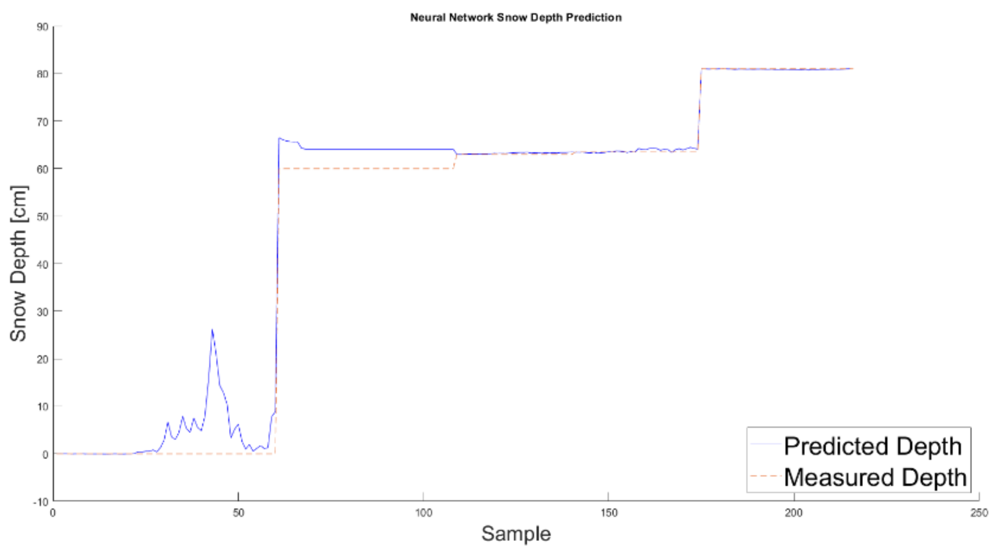

Figure 8: The model for estimating the snow depth based on the BR algorithm in MATLAB.

Gaussian Process Regression (GPP). These models have a challenge in estimating the depth and density in November with no snow. A new model was developed based on neural network with one hidden layer of 50 neurons, using Bayesian Regularization (BR). Figure 8 shows the measured depth and the predicted depth based on the BR model.

The model seems to predict the snow depth ok except for the period in November with no snow, and better than the GPP model. The corresponding model for measured density and predicted density is shown in Figure 7.

The model seems to predict the snow density ok, even the November is ok even if there is no snow. There is also a challenge with the density in February, this will be part of the discussion section.

4.2 Model trained using ML.NET

ML.NET also contains an automatic trainer but only the algorithm with the best validation results regarding MSE is available at the end of the development process. A test application in C# was developed to plot the measured and predicted values from the models. None of algorithms for the automatic trainer made a model as good as the BR model from MATLAB, and the best results was from the Fast Forrest algorithm. The snow density result is shown in Figure 9 and the snow depth result is shown in Figure 10.

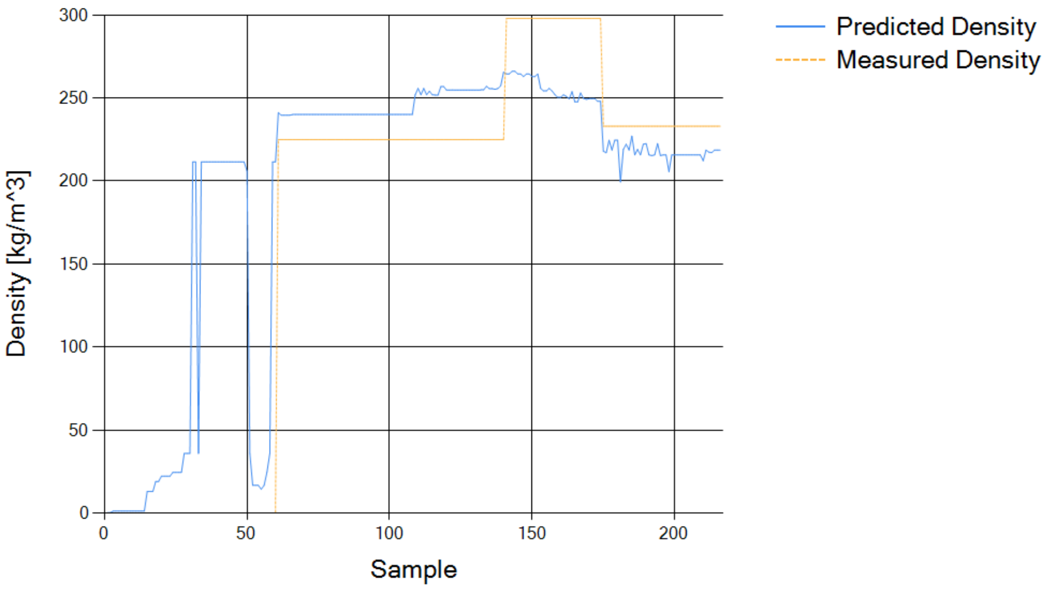

Figure 9: The model for estimation of the snow density based on the ML.NET algorithm.

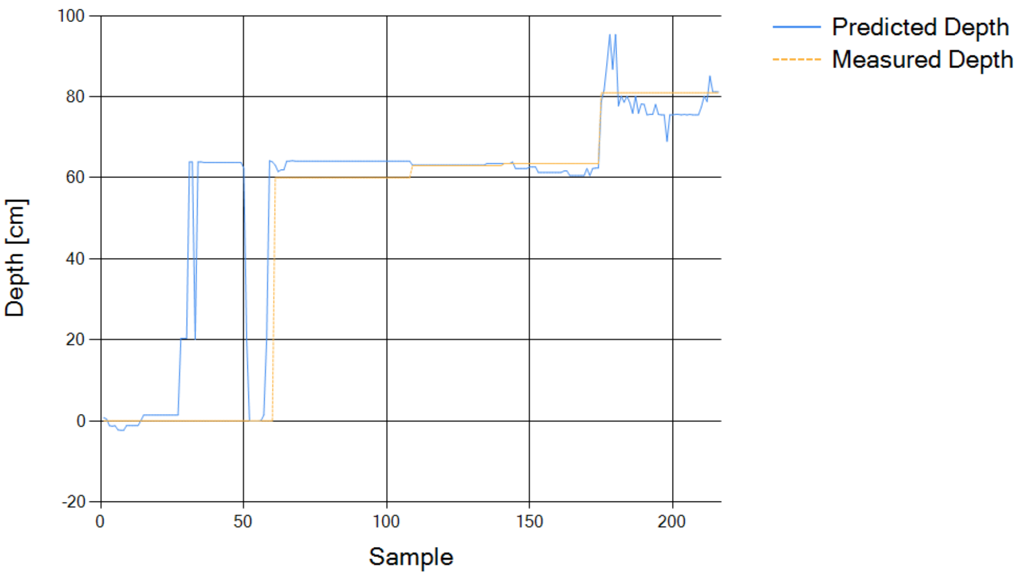

Figure 10: The model for estimation of the snow depth based on the ML.NET algorithm.

Figure 10 shows that the model will not estimate the snow depth very good in November when there is no snow. Most of the models seems to have a challenge with this period. The same challenge with the snow density estimation for the model shown in Figure 9. The models developed in the MATLAB framework in Figure 7 and Figure 8 seem to estimate the snow depth and snow density better then the models developed in the ML.NET framework.

4.3 Model trained using Keras

Keras is a Python based application programming interface (API) for TensorFlow. Neural network models and the Adam optimizer was used in Keras, with 80% of the samples for training and 20% of the samples used for validation. The selection of samples for these sets were randomized and 1000 epochs was the default for training the models.

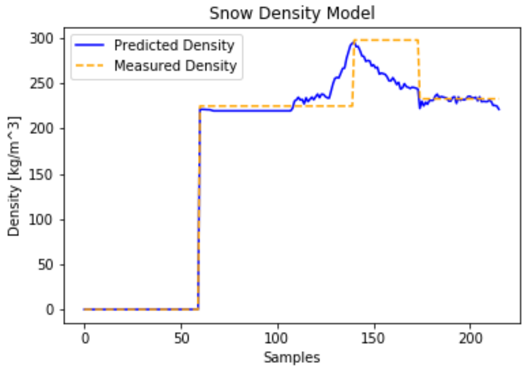

Figure 11: The snow density prediction and measurement based on the Keras framework.

The snow depth model architecture consists of eight hidden layers with between 50 and 200 neurons in each hidden layer. The activation functions used in these layers are Tanh, Relu and Sigmoid types.

The measured and estimated snow depth, based on the Keras model, is shown in Figure 12. The model shows a good prediction also for the November period when there was no snow.

The snow density model architecture consists of six hidden layers with between 25 and 150 neurons in each hidden layer. The activation functions used in these layers are Tanh and Relu types. The architecture for the snow depth model is more complex than the snow density model.

The measured and estimated snow density, based on the Keras model, is shown in Figure 11. The snow density model shows a good prediction of the snow denisty although the density at higher levels is not following the reference optimally.

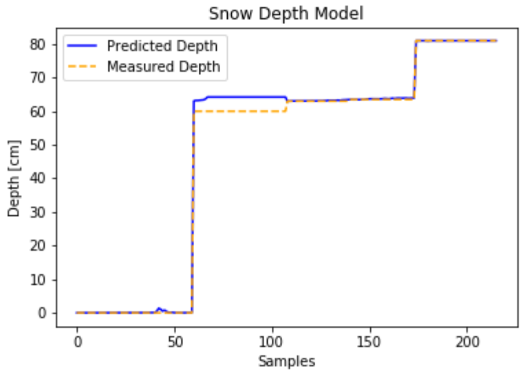

Figure 12: The snow depth prediction and measurement based on the Keras framework.

Since both the snow depth and the snow density models based on the Keras framework seems to perform best on the limited data set, these models were selected as the input for the SWE model.

The SWE model, shown in Figure 13, showing a good prediction of the SWE. Some deviations in the upper regions but still a good fit.

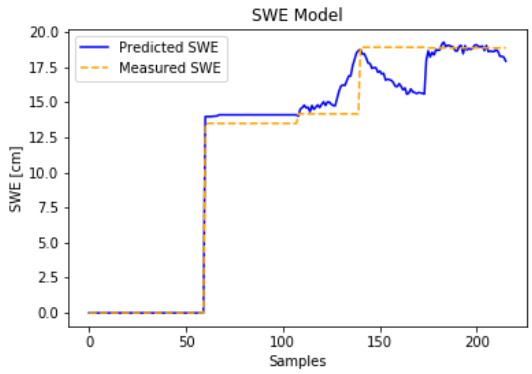

Figure 13: The snow water equivalent (SWE) prediction and measurement based on the Keras framework.

5 Discussion

There are several comments that should be given based on these measurements and results. Many of the models seems to have a challenge predicting the correct snow depth and snow density when no snow, so other weather or environment parameters should be checked and possible added to the system. As always with machine learning methods more data is wanted. In this case there are too few reference measurements with too large jumps in snow depth. This time the reference was based on manual intermittent measurements. The depth measurements were ok, while the density was more of a challenge. First the density was assumed to be almost constant during a day, but in a sunny day the density could vary a lot. The next step will be to have an automatic measurement of the depth and the density as well. Some experiments have been done using an ultrasonic sensor during the winter and if the calculation is compensated for the temperature and humidity, an ultrasonic sensor will work as depth measurement. A weight cell will also be considered to have a reference measurement for the weight of the snow, estimate the density based on the weight and the depth.

A big challenge with the capacitance sensor devices is that these devices are contact sensors and during the springtime when the snow is melting, the snow crystals will change the shape and size, and will lose contact with fixed objects (Muller, 2020). The capacitance sensor will work in wintertime with temperatures below 0ºC but will lose contact with the snow in springtime and temperatures above 0ºC. This is the reason the last dataset, the dataset from April was not used in the model development process. The current design is not a good solution, another solution should be used to measure the capacitance in the next version of the prototype.

The first versions of the models estimate the parameters as expected. However, the amount of labeled data was too small to get a good and reliable model, and the number of capacitive sensors gave a too rough estimate of the snow depth.

The test and validation set are both collected from the same original dataset, with small variations, especially in the reference. It is therefore likely that the presented Keras model with a huge number of hidden neurons are partly overfitted. Additional new datasets should therefore be tested before the models can be assessed for more conditions in a broader range.

Future work will involve a more robust logging system, a non-contact depth sensor like ultra-sonic or laser sensor and a weight sensor for better calibration data. A new design of the capacitance sensor is needed to better measure the capacitance of the snow during the springtime.

The measurement system is measuring the capacitance at different layers in the snowpack, and estimating the snow depth, the density, and the water content. The focus in this project has been on measurements for hydro power systems. However, properties for other systems like for skiing or avalanches should also be possible with this type of measurements. Skiing properties like snow depth and surface conditions can be estimated based on the measured values from the capacitance sensors. Avalanches are beyond the scope of this study. However, by measuring the capacitance in many layers in the snowpack, with a higher resolution of sensor devices, estimating the condition of each layer can be used for evaluating the risk of avalanches in that area.

6 Conclusion

Three different machine learning framework was used, MATLAB, ML.NET and Keras. The models developed using the Keras framework were the best models, especially in the period with no snow. Two different models were developed for each framework, for prediction of the snow depth and snow density. All models performed ok with a limited set of data samples. The snow density and snow depth models from the Keras framework was used for the SWE model, the goal of this work. The model predicting the SWE was working ok with some limitation. However, the limited set of samples was the largest limitation of making a good model.

References

J. M. Bjerke, M. N. Murillo Abril, A. Jaganjac and N. Pouladi, *Measurement of snow density*, “MSc. Project”, University of South-Eastern Norway (USN), Porsgrunn, 2019.

Ccin, Current Snow Cover, Snow Anomaly Tracking, https://ccin.ca/ccw/snow/current/swe, Canadian Cryospheric Information Network, University of Waterloo, 2021.

A. Denoth, A. Foglar, P. Weiland, C. Mätzler, H. Aebischer, M. Tiuri and A. Sihvola, A comparative study of instruments for measuring the liquid water content of snow, *Journal of Applied Physics*, 1984, DOI: 10.1063/1.334215.

A. Denoth and I. Wilhelmy, Snow dielectric devices and field applications, International Snow Science Workshop, 1988.

M. Hallikainen, F. Ulaby and M. Abdel-Razik, Measurements of the dielectric properties of snow in the 4-18 GHz frequency range, *12th European Conference on Microwave*, 1984.

K. Muller, H. Toft Larsen, and G. Sojer, Snøomvandling (in Norwegian), Fact sheet, NVE, 2020

M. N. Murillo Abril, *Development of a remote measurement node for snow density*, “MSc. Thesis”, University of South-Eastern Norway (USN), Porsgrunn, 2020.

M. N. Murillo Abril, B. Furenes, N.-O. Skeie, Development of a model to estimate parameters in a snowpack based on capacitive measurements, *SIMS conference*, 2020.

M. Niang, M. Bernier, M. Stacheder, A. Brandelik, and E. van Bochove, Influence of snow temperature interpolation algorithm and dielectric mixing-model coefficient on density and liquid water content determination in a cold seasonal snowpack, *Sensing and Imaging*, 2006, DOI: 10.1007/s11220-006-0020-9.

H. B. Stranden, B. Lirhus Ree and K. M. Møen, Recommendations for automatic measurements of snow water equivalent in NVE, Norwegian Water Resources and Energy Directorate (NVE), 2015.

F. Techel and C. Pielmeier, Point observations of liquid water content in wet snow – investigating methodical, spatial and temporal aspects, *The Cryosphere*, 2011.

H. N. Vahl, *Development of models for estimating snow depth and snow density using machine learning methods*, “MSc. Thesis”, University of South-Eastern Norway (USN), Porsgrunn, 2021.

Wsl, Long-term snow water equivalent measurements, https://www.wsl.ch/en/projects/swe-measurements.html, WSL Institute for Snow and Avalanche Research SLF, 2021.