



## DEPARTMENT OF MATHEMATICAL SCIENCES

MA3911 - MASTER THESIS IN MATHEMATICAL  
SCIENCES

---

# Site-specific probabilistic forecast for avalanches

---

*Author:*

Benjamin Sigbjørnsen

*Supervisor:*

Ingelin Steinsland

Date

1st June 2024

---

## Acknowledgements

I want to thank my supervisor, Ingelin Steinsland, for a great collaboration and for her invaluable help and guidance. I also wish to express my gratitude to Statens Vegvesen and particularly Eivind Schnell Juvik for their essential assistance in data collection and for providing insight and information about avalanches and avalanche forecasting.

---

# Abstract

The aim of this research is to establish and validate a statistical method for a site-specific probabilistic forecast for avalanche releases. A probabilistic forecast is a forecast combining both predictions and uncertainty of predictions. To be able to quantify the uncertainty of predictions is important for local avalanche forecasting. An avalanche hitting an open road could endanger the life of people in the area. A closed road can have economical and human consequences.

Avalanche forecasting in Norway today is a manual warning issued as a danger level by an expert forecaster based on observations of the snow quality and the weather, weather forecasts, and more. There has been attempts to automate avalanche warnings. Earlier research has indicated that different machine learning algorithms can be used to forecast avalanches.

In this thesis, different statistical and machine learning methods are suggested and validated for a case study. The case study chosen is two winter seasons from the avalanche-prone road stretch, Holmbuktura in Troms County, in Northern Norway. The reason for choosing Holmbuktura is the existence of a radar detecting thermal activity, which has multiple years of detection data on avalanche releases in several avalanche paths. Weather forecasts and weather observations available the day before the prediction are used as explanatory variables. In addition, daily avalanche count from satellite detection from a nearby area are included as explanatory variables. Our goal is to establish methods for probabilistic one-day-ahead daily avalanche count forecasts based on these explanatory variables. The methods suggested and explored are different types of generalized linear models, tree-based methods, and the method of nearest neighbours. Ensemble methods combining different forecasts are also explored. To validate the suggested method we calculate mean continuous ranked probability score and root mean squared error based on cross-validation and do visual inspection based on PIT-diagrams. The resulting models are compared to two baseline models; (1) A generalized linear model with danger level as the only explanatory variable and (2) a generalized linear model without explanatory variables.

The analysis found that probabilistic ensemble models was best at forecasting avalanche count. It performs better than both baseline models. The predictions from the random forest are the most important input for the ensemble models.

---

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Background</b>	<b>4</b>
2.1	Avalanches . . . . .	4
2.2	Probabilistic forecasting . . . . .	6
2.2.1	Introduction . . . . .	6
2.2.2	Evaluating probabilistic forecast . . . . .	7
2.3	Forecasting methods . . . . .	9
2.3.1	Tree analysis . . . . .	9
2.3.2	The method of nearest neighbours . . . . .	11
2.3.3	Generalized linear models . . . . .	11
2.3.4	Ensemble forecasting . . . . .	12
<b>3</b>	<b>Location and data collection</b>	<b>13</b>
3.1	Holmbuktura . . . . .	13
3.2	Case study and data . . . . .	15
3.2.1	Avalanche data from Holmbuktura . . . . .	15
3.2.2	Avalanche data from satellites . . . . .	16
3.2.3	Weather forecast . . . . .	18
3.2.4	Observed weather . . . . .	23
3.2.5	Danger level . . . . .	23
3.2.6	Correlation analysis between types of data . . . . .	24
3.3	Choosing reduced set of explanatory variables . . . . .	24
<b>4</b>	<b>Modelling</b>	<b>26</b>
4.1	Sets of explanatory variables . . . . .	26
4.2	Tree models . . . . .	28

---

---

4.2.1	The models . . . . .	28
4.2.2	Evaluating the models . . . . .	28
4.3	Models fitted using method of nearest neighbours . . . . .	29
4.3.1	The models . . . . .	29
4.3.2	Evaluating the models . . . . .	30
4.4	Generalized linear models . . . . .	30
4.4.1	The models . . . . .	30
4.4.2	Evaluating the models . . . . .	31
4.5	Ensemble model . . . . .	31
4.5.1	The model . . . . .	31
4.5.2	Evaluating the model . . . . .	32
4.6	Evaluating models on example dates . . . . .	32
<b>5</b>	<b>Results</b>	<b>33</b>
5.1	Tree models . . . . .	33
5.1.1	Simple regression trees . . . . .	33
5.1.2	Random forests and boosting models . . . . .	36
5.1.3	Performance of tree-based models . . . . .	38
5.2	Models fitted using the method of nearest neighbours . . . . .	38
5.3	Generalized linear models . . . . .	39
5.4	Ensemble model . . . . .	42
5.4.1	Choosing ensemble model . . . . .	42
5.4.2	Ensemble model fit . . . . .	42
5.5	Ensemble model with random forest prediction as the only explanatory variable . . . . .	43
5.6	Comparison of the different model types . . . . .	45
5.6.1	Using example dates . . . . .	45
5.6.2	Comparing models with baseline models . . . . .	47

---

---

<b>6 Discussion and conclusion</b>	<b>48</b>
6.1 Discussion . . . . .	48
<b>Bibliography</b>	<b>50</b>
<b>Appendix</b>	<b>52</b>
A Data exploration . . . . .	52
B Tree models . . . . .	54
C Models fitted using the method of nearest neighbours . . . . .	55
D Generalized linear models . . . . .	56
E Code . . . . .	59

---

# 1 Introduction

Avalanches are defined as a natural phenomenon where snow masses are moving down a steep hillside (Schweizer et al., 2003, p. 1). In Norway, avalanches inflict a substantial amount of damage. There have been 1510 registered deaths related to avalanches in Norway between 1836 and today (Lied and Kristensen, 2003, p. 13). The traffic is also heavily disrupted by avalanches. Annually, is expected that 10 to 15 vehicles get taken by avalanches and that around 150 roads get closed each winter in Norway due to avalanches (Lied and Kristensen, 2003, p. 11). The quality of avalanche risk forecasts is therefore important for the reliability and safety of roads in areas prone to avalanches. A closed road impacts the traffic and the lives of the people using it. An accurate forecast will therefore close the road when and only when necessary. The other way to make a road stretch prone to avalanches safe is to build a tunnel. However, a tunnel is expensive to build. A tunnel is expected to cost 230 000 NOK per meter. This is more than three times more than the price for an ordinary road (Eivind Schnell Juvik, personal communication, October 2023). The avalanche forecast in Norway today is a regional warning based on expert opinion, described in Section 2.1.

The case study selected is Holmbuktura. The relationship between daily intensity of avalanche releases in Holmbuktura and the various explanatory variables are considered. The avalanches hitting the road in Holmbuktura are the ones of interest, but all avalanches detected in an specific area are included in the analysis to get more data. Holmbuktura is located in Troms county in Norway. Holmbuktura is chosen in cooperation with Eivind Schnell Juvik, Geotechnical Engineer/Avalanche Expert Advisor in Statens Vegvesen. This road is selected because it is in an avalanche-prone area with a lot of avalanche releases in the hillside. However, the most important reason for choosing this area is the existence of a radar recording avalanches. The radars can detect avalanches in the mountainside with high precision. During a test run of the radar in Holmbuktura in 2017 and 2018, it was found that all the relevant avalanche releases was detected, and that the false alarm rate was 3.4 percent (Meier, 2018, p. 16). This means the data from the radar is reliable. It is also important the avalanches are naturally released. More information about Holmbuktura can be found in Section 3.1. The methodology used in Holmbuktura could be used for other road stretches prone to avalanches. The radar automatically closes the road when an avalanche is detected. The radar detects avalanches in an area in Holmbuktura where the avalanches potentially could hit the road. The radar also opens the road if the avalanche does not end up hitting the road. Figure 1 depicts a sign containing information about where the radar is located and how the radar works.



Figure 1: Information sign standing in Holmbuktura illustrating how the radar works.

Similar research has been done by fitting random forests and by doing logistic regression on avalanche data from Senja (Hennum, 2016). Hennum (2016) found that the random forest was best able to predict avalanche activity. The advantage of regression with a Poisson distributed response instead of a logistic response is that there is more information used when taking into account the avalanche count. Buser (1983) has looked into the use of the method of nearest neighbours to predict avalanche releases. Davis et al. (1999) has used tree analysis to look at relationship

---

between weather variables and avalanches. Gneiting and Katzfuss (2014) proposes ensemble forecasting as a way to do probabilistic forecasting. Ensemble forecasts as weather forecasts has been implemented with great success (Gneiting and Katzfuss, 2014, p. 140), and has also become important in other fields (Gneiting and Katzfuss, 2014, p. 143). Data-driven avalanche forecasting using machine learning algorithms is a topic under research. Both Herwijnen et al. (2023) and Viallon-Galinier et al. (2023) have conducted recent studies on this topic.

The aim of the thesis is to establish and validate a statistical method for site-specific probabilistic forecast for avalanche releases. A site-specific probabilistic forecast for avalanches could be a useful tool for operational avalanche forecasting when deciding whether a road should be closed or not. Various statistical models with different sets of explanatory variables will be tested, both to do statistical inference about avalanche risk and to predict the risk. Data from weather forecasts, radar detecting avalanches, and satellites from Statens Vegvesen (the Norwegian Public Roads Administration) detecting avalanches are considered as possible explanatory variables in this thesis. The satellite data is a new type of data. Statens Vegvesen has invested in acquiring and analyzing satellite data to detect avalanches. Therefore, it is interesting to evaluate the value of this satellite data in relation to avalanches. Different tree models, models fitted using the method nearest neighbours, generalized linear models, and an ensemble model are fitted to the data. The generalized linear models and the ensemble model are probabilistic models, while the machine learning approaches with tree-based models and the models fitted using the method of nearest neighbours are deterministic.

The thesis has some relevance for sustainability with respect to the United Nations' sustainability goals. Goal 9 of the United Nations' sustainability goals is about infrastructure (United Nations Department of Economic and Social Affairs, n.d.-a). Target 9.1 in this goal is: "Develop quality, reliable, sustainable and resilient infrastructure, including regional and transborder infrastructure, to support economic development and human well-being, with a focus on affordable and equitable access for all" (United Nations Department of Economic and Social Affairs, n.d.-b), which corresponds with the purpose of this task. Better avalanche predictions will provide safer and more reliable roads. Goal 13 of United Nations' sustainability goals is to stop climate changes (United Nations Department of Economic and Social Affairs, n.d.-a). A tunnel requires a lot of cement, and cement industry is a one of the biggest emitters of carbon dioxide and thus a contributor to global warming. The cement industry is responsible for 8 percent of the total annually global emissions (Ellis et al., 2020). Hence, it is preferred to avoid building a tunnel if possible to avoid carbon dioxide emission. Better avalanche predictions could therefore contribute towards reaching this goal.

In the next section, Section 2, the theory relevant for understanding the approach and results of the thesis are introduced. In Section 3 the study area Holmbuktura and the data collected are described. Further, in Section 4 the proposed models used to predict number of avalanches are explained, and in Section 5 the results from fitting the models to the data are presented. In the last Section, Section 6, the results are discussed, and conclusions are drawn from the findings.

---

## 2 Background

### 2.1 Avalanches



Figure 2: An avalanche is released in the release zone (utløsningsområdet) in the top area, continues into the avalanche track (skredbane) in the middle area, and then ends in the runout zone (utløpsområdet) in the bottom area.

Source: Lied and Kristensen, 2003, p. 21



Figure 3: Avalanche hitting the road in Holmbuktura in 2012.

Source: From news article in NRK <https://www.nrk.no/tromsogfinnmark/over-100-personer-isolert-av-skred-1.8036654> Retrieved 28.01.2024.

There are different types of avalanches, and the two main categories are slab avalanches and loose snow avalanches (Schweizer et al., 2003, p. 1). Slab avalanches are released by a big slab starting to move, while the loose snow avalanches are released from a point in unstable snow masses (Schweizer et al., 2003, p. 1). Slab avalanches and loose snow avalanches can occur under various conditions (Varsom, n.d.-b): Fresh snow can lead to loose snow avalanches if the temperature is high and if there is little wind. Slab avalanches can also be released after a snowfall if there are wind or high temperature that make the fresh snow form slabs. Wet snow from rain or high temperatures can cause both slab and loose snow avalanches. Furthermore, wind-drifted snow from wind in combinations with fresh snow can lead to loose snow avalanches.

Information about where avalanches can hit is described by Varsom (n.d.-b). If the slope is 30 degrees or more, avalanches can be released. Avalanches are rarely released when the incline is less than 30 degrees, but can continue into terrain with less steep slopes, called runout zones. Release zones together with runout zones form avalanche terrain, terrain where avalanches could hit. Figure 2 shows a typical avalanche terrain. Avalanches are released in the hillside and can continue into more level terrain for some time until they stop.

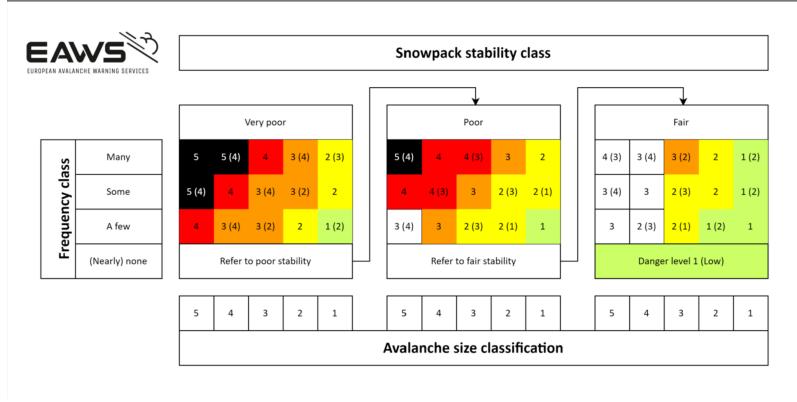


Figure 4: Avalanche matrix used to set danger level for avalanches.

Source: European avalanche warning services, n.d.

In Norway today, avalanche warnings are issued as a manually set danger level on a scale from 1 to 5 (Varsom, n.d.-a). The danger level is determined by expected avalanche size, expected stability of snowpack, and expected frequency of avalanches, and these factors are combined into a matrix called ADAM matrix in Figure 4. The matrix in Figure 4 is the matrix valid for the winter from 2023 to 2024. The matrix are used to standardize the warnings. The danger level is set manually using snow and weather observations and forecasts, and the goal of the matrix is therefore to make the danger level independent of the warning issuer. The warning issuer follows a guide when issuing the warning, and setting the danger level (Eivind Schnell Juvik, personal communication, November 2023). The warnings are issued on a regional level. The highest expected danger level for an area with a size of minimum 100 squared kilometers sets the danger level for the whole region. The warnings are valid for a day.

## 2.2 Probabilistic forecasting

### 2.2.1 Introduction

Forecasting using predictions and probability distributions of the predictions is called probabilistic forecasting (Gneiting and Katzfuss, 2014, p. 126). Our goal is to predict avalanche count hitting the road, and therefore we consider probability mass functions, where the following conditions have to hold:

$$\sum_Y f(Y_i|x) = 1 \quad (1)$$

$$f(Y_i|x) \geq 0 \quad \forall y_i \quad (2)$$

$Y_i$  are possible discrete outcomes and  $x$  is the vector of explanatory variables. The probability distribution of a forecast is of interest when it is desirable to quantify the uncertainty of the forecast. Probabilistic forecasting has in the last decades become more and more popular, and has been used in many different types of

---

forecasting, including flood risk assessment, seismic hazard and election outcomes (Gneiting and Katzfuss, 2014, p. 126). When an avalanche is released is hard to predict accurately, and hence these forecasts should also be probabilistic. Avalanche releases are difficult to predict, but they depend on multiple factors like weather, snow quality, and terrain. This is elaborated in Section 2.1 about avalanches.

### 2.2.2 Evaluating probabilistic forecast

To assess the forecast, we have to look at the calibration and sharpness of the predictions. A calibrated prediction of avalanche count would have unbiased expectation and variance that matches with observed variance. The sharpness of a prediction refers to the uncertainty of the predictions and the magnitude of the uncertainty.

The Probability Integral Transform (PIT) is used to check if the calibration of the proposed model matches with the real-world realizations (Haben et al., 2023, p. 106). It is defined as the random variable

$$Z_F = F(y_{obs}) \quad (3)$$

, where  $F$  is a cumulative mass function, and  $y_{obs}$  is an observation of the event that the model want to forecast (Gneiting and Katzfuss, 2014, p. 130). A probabilistic forecast can be evaluated by looking at the distribution of  $Z_F$  for the different observations. The forecast is considered to be calibrated and dispersed correctly if the distribution of the random variable  $Z_F$  is standard uniform. The uniformity of the distribution of  $Z_F$  can be assessed by looking at the histogram of the variables. In the discrete case, the histogram of PITs would not be uniform. There exists several techniques to transform the PITs such that the distribution of the random variables are uniform (Czado et al., 2009, p. 2). If the distribution are not uniform, then there might be a problem with the variance. An overdispersed model, a model with too large variance, will have a concave-looking histogram. On the other side, an underdispersed distribution, a model with too small variance, will have a convex looking histogram of PITs.

The following paragraph is based on Gneiting and Raftery (2007, p. 359-361). To assess the quality of the forecast, strictly proper scoring rules can also be used. Scoring rules evaluates the forecast by giving numerical scores to the realizations based on the proposed distribution of the forecast. The numerical scores are denoted  $S(P, y_{obs})$ , where  $P$  is the proposed probabilistic forecast, and  $y_{obs}$  is a realization of an event.  $S(P, Q)$  is defined as the expected proper scoring rule, where  $Q$  is the the best guess of the distributional forecast. A scoring rule is said to be proper if  $S(Q, Q) \geq S(P, Q)$  and strictly proper if there is equality only when  $P = Q$ . Different forecast are compared by looking at the average scores

$$S_n = \frac{1}{n} \sum_{i=1}^n S(P_i, y_{obs,i}) \quad (4)$$

of  $n$  observation.

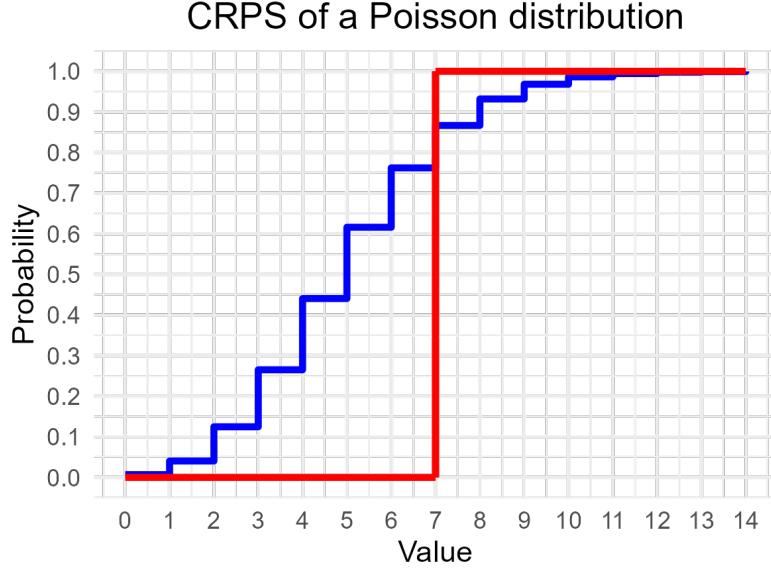


Figure 5: Plot of cumulative distribution function of a Poisson model with expectation 5 and an indicator function  $\mathbb{I}\{y \geq 7\}$ .

Mean continuous ranked probability score (CRPS) are used to check both the sharpness and calibration of the models. The mathematical definition of continuous ranked probability score is stated in (5):

$$CRPS(F, y_{obs}) = \int_{-\infty}^{\infty} (F(y) - \mathbb{I}\{y \geq y_{obs}\})^2 dy, \quad (5)$$

where  $F(y)$  is the cumulative distribution function of the probabilistic forecast  $f(y)$  and  $\mathbb{I}\{y \geq y_{obs}\}$  is an indicator function which is equal to 1 if the condition  $y \geq y_{obs}$  is true and 0 otherwise (Gneiting and Raftery, 2007, p. 367). In Figure 5 an example cumulative distribution function and indicator function is plotted. The area between the functions corresponds to the Continuous ranked probability score. CRPS is used when it is desirable to reward models that assigns high probability to values close to the true value, and when both deterministic and probabilistic forecasts are evaluated. CRPS is a strictly proper scoring rule for the cases that this thesis covers (Gneiting and Raftery, 2007, p. 367).

The root mean squared cross-validation error (RMSE) can also be used to check the sharpness of the models. The definition of RMSE is stated in (6).

$$RMSE = \sqrt{\frac{1}{k} \sum_{j=1}^k MSE_j}, \quad (6)$$

where  $MSE_j$  is the mean squared error of the  $k$ -th fold in  $k$ -fold cross-validation. Performing  $k$ -fold cross-validation divides the data into  $k$  folds, where  $k-1$  of the folds are used for training the model, while the last one is used for evaluation.

Mean squared error is defined in (7).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_{obs,i})^2 \quad (7)$$

## 2.3 Forecasting methods

In this section, the methods and model cases in use for avalanche forecasting in this thesis are introduced. The output of the models is denoted  $f(x)$ . For the probabilistic models,  $f(x)$  will be a probability density function, while for the deterministic functions it will be a numeric response.

### 2.3.1 Tree analysis

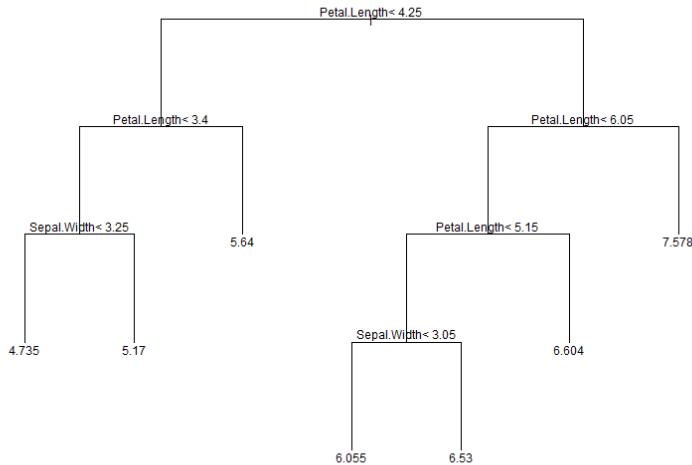


Figure 6: Example regression tree fitting using iris flower data with sepal length as response. Iris flower data is often used as example data set in R.

Tree models divide the sample space of the explanatory variables into sections with different responses.(James et al., 2021, p. 327). The sample space are divided iteratively in a way that reduces the error the most until a stop criteria is reached. (8) shows how the response of a tree model,  $f(x)$ , can be expressed mathematically. The  $i$ -th section is denoted  $A_i$ , while the  $i$ -th response are denoted  $C_i$ .  $N$  is the total number of partitions.

$$f(x) = \begin{cases} C_1 & \text{if } x \in A_1 \\ C_2 & \text{if } x \in A_2 \\ \vdots \\ C_N & \text{if } x \in A_N \end{cases} \quad (8)$$

---

Figure 6 illustrates a regression tree fitted using iris flower data with sepal length as response. A regression tree models a continuous response. Further, it can also be seen in Figure 6 that the tree consists of nodes, splitting points of the tree, and branches that are lines between nodes. The terminal nodes have responses attached. If the inequality are true, the left branch of the tree is taken to next node, otherwise the right node are taken. This is iterated from the top node, the root, until a terminal node is reached. Pruning is a way of avoiding overfitting by reducing the number of terminal nodes to reduce variance and increase the interpretability, at the cost of higher bias (James et al., 2021, p. 331). The number of terminal nodes can be chosen by cross-validation.

Boosting and random forests are techniques to improve the accuracy of the predictions of trees (James et al., 2021, p. 340). Random forests are built by growing multiple trees from bootstrapped data, and then averaging the responses as seen in (9)

$$f(x) = \frac{1}{N} \sum_{i=1}^N f_i(x), \quad (9)$$

where  $f_i(x)$  are tree number i of the total N trees. (James et al., 2021, p. 344). At each splitting point of the tree, a random subset of explanatory variables is chosen. These subsets are of a predetermined size. When employing bootstrapping and averaging in this manner, variance is reduced, leading to improved predictions. However, this improvement comes at the cost of reduced interpretability. This next part about boosting is based on James et al. (2021, p. 347). In boosting, multiple trees are also grown. These trees are grown sequentially. The first tree,  $f_1(x)$  are fitted from the observations,  $y$ . The next trees are fitted from the residuals defined in (10) Tree j are fitted from the residuals,  $r$ , after fitting the j-1 previous trees.

$$r = y - \sum_{k=1}^{j-1} \lambda f_k(x) \quad (10)$$

The response is a sum of the n trees fitted. The trees fitted are called  $f_i(x)$ , where i is an index ranging from 1 to n and x is the vector of explanatory variables. This is shown in (11).

$$f(x) = \sum_{i=1}^n \lambda f_i(x) \quad (11)$$

$\lambda$  in (11) is called the shrinkage parameter. It controls the learning speed of boosting. As with random forests, interpretability are lost, but predictions are often improved, while variance is reduced.

---

### 2.3.2 The method of nearest neighbours

The idea behind the method of nearest neighbours is that observations with similar explanatory variables will have similar response (Hastie, 2009, p. 14). The parameters that are similar should be parameters that are actually affecting the response. To analyze the response using nearest neighbours, data on the responses and explanatory variables, expressed as vectors of explanatory variables related to the response, are needed. Further, it is necessary to calculate distance between the vector of explanatory variables of the date that are of interest to predict, and the observed vectors of explanatory variables.

For this to work, the columns in the design matrix has to be scaled so that they have the same mean and standard deviation. Without this standardization, the distance between some of the explanatory variables could be more important than others.

$$d(x_0, x_i) = \|x_0 - x_i\| \quad (12)$$

(12) illustrates how the distances between the vectors of explanatory variables of the nearest neighbours,  $x_i$ , and the vector of explanatory variables of the response of interest to predict  $x_0$  are calculated. The responses and the vectors of explanatory variables with the smallest distances to the vector of explanatory variables of interest are then chosen. The responses are then combined into a predicted response. The number of responses, the number of neighbours, in use to predict the response of interest must be chosen.

$$f(x_0) = \frac{1}{K} \sum_{i=1}^K N_i \quad (13)$$

In (13), the method of nearest neighbours are presented. There are  $K$  neighbours, and  $N_i$  is the response of one of the  $K$  nearest neighbours. If the 4 closest vectors of explanatory variables have the responses 2,2,4, and 4, then the predicted response would be  $\frac{1}{4}(2 + 2 + 4 + 4) = 3$ .

An important thing to note about the method of nearest neighbours is that predictions done on training data should have an error function approximately increasing with the number of neighbours, and an error function being equal to zero when the number of neighbours are 1 (Hastie, 2009, p. 15).

### 2.3.3 Generalized linear models

Generalized linear models are models that links a linear combination of explanatory variables with a response (Fahrmeir et al., 2013, p. 269). The normal distributed linear models are generalized. If the response is in the exponential family, the

---

probability distribution function of the univariate response can be written as:

$$f(x|\theta) = \exp\left(\frac{x\theta - b(\theta)}{\phi}w + c(x, \phi, w)\right), \quad (14)$$

where  $b(\theta)$  and  $c(\theta)$  are functions,  $\theta$  is the parameter of interest, and  $w$  and  $\phi$  are constants (Fahrmeir et al., 2013, p. 301). A generalized linear model has the following properties:

1. A random response  $\mu$ .
2. A systematic component  $\eta_i = x_i^T \beta$ , where  $x_i^T$  are the i-th row in the design matrix,  $X$ , and  $\beta$  are the regression parameters.
3. A link function  $g(\mu) = \eta_i$ .

The Poisson distribution belongs to the exponential family, and it is therefore possible to model the response of a generalized linear model as a Poisson distributed variable. This is usually done through a link function  $g(\lambda_i) = \log(\lambda_i) = \eta_i$ , which is the canonical link of a generalized linear model with a Poisson distributed response. The model parameters,  $\beta$ , are fitted by minimizing the error of equation (15).

$$\ln(Y) = \beta X + \epsilon, \quad (15)$$

where  $Y$  is a column vector of observations,  $X$  is a design matrix containing explanatory variables, and  $\epsilon$  is the error.

The Poisson distribution has the probability distribution shown in (16).

$$f(x|\lambda) = \frac{\lambda^x}{x!} \exp(-\lambda) \quad (16)$$

This is a discrete distribution used for counting processes.

### 2.3.4 Ensemble forecasting

Ensemble forecasting is a type of forecasting where the forecast is issued by combining other point forecasts (Gneiting and Katzfuss, 2014, p. 140). The forecast of the ensemble forecast could however be probabilistic, either by using the variance of the forecasts included as explanatory variables, or by assuming the response has some distribution. Ensemble forecasting has been used for weather forecasts by combining different forecasts with different initial conditions (Gneiting and Katzfuss, 2014, p. 140).

---

### 3 Location and data collection

#### 3.1 Holmbuktura

Figure 7 shows Holmbuktura being located in the northern part of Norway. Holmbuktura is situated to the southeast of the city Tromsø, in a region called Lyngen, as depicted in Figure 8. There is a fjord on the west side of Holmbuktura, and tall mountains on the east side, meaning the hillside is facing west. The topography is illustrated in the map in Figure 10. There is also a road running in between the mountains and the fjord. Figure 9 provides a picture of the path of the road between the mountains and the water, seen from northwest. Holmbuktura is located at the 69th latitude.



Figure 7: Geographical position of Holmbuktura in Norway.

Source: Google Maps, 10th February 2024, <https://www.google.com/maps>



Figure 8: Geographical position of Holmbuktura in Troms. The straight-line distance between Tromsø and Holmbuktura is 35 kilometer.s

Source: Google Maps, 30th May 2024, <https://www.google.com/maps>



Figure 9: Holmbuktura seen from the road located to its northwest.

Source: Google Street View, 8th November 2023, <https://www.google.com/maps>



Figure 10: Topological map of Holmbuktura showing height curves and some heights.

Source: Kartverket, 1st February 2023, <https://hoydedata.no/LaserInnsyn2/>

Holmbuktura has been chosen because it is an avalanche-prone area. The total number of avalanche releases in Holmbuktura used in the analysis is 115. Figure 9 and 10 shows Holmbuktura having really steep slopes. This is a necessity for avalanches to be released, as discussed in Section 2.1, and the steep slopes are therefore one of the reasons why Holmbuktura is an avalanche-prone area. The geographical location of Holmbuktura also makes the area prone to avalanches. The northern location of Holmbuktura in Norway results in snowy winters, and the fresh snow increases the avalanche risk as discussed in Section 2.1. The road stretch is in the runout zone of the avalanches. Figure 3 shows an avalanche hitting the road in Holmbuktura.

### 3.2 Case study and data

Data from the winter seasons of 2019-2020 and 2020-2021 in Holmbuktura is used in the analysis. The winter season is defined as lasting from 1st of December to 31st of May. Both winter seasons are used for model fitting.

#### 3.2.1 Avalanche data from Holmbuktura

The avalanche data from the radar in Holmbuktura was provided by Wyssen Norge. Wyssen Norge is responsible for the avalanche-detecting radar in Holmbuktura (Amundsen, 2019). According to geologist Andreas Persson in Statens Vegvesen, the radar uses the Doppler effect to detect the release of avalanches (Amundsen, 2019).

---

In the dataset, there is a row for every registered avalanche with a start and end time of the avalanche release, and information about the avalanche size and speed. Information about whether the avalanche hit the road or not are also included in the dataset. The number of avalanches in one day and information about whether the avalanche hit the road or not are used in the data analysis. Table 1 presents a frequency table of avalanche count per day for two winter seasons. Time series plot of the data can be found in Figure A.1 and A.2 in Appendix A.

Table 1: Frequency table of radar observations of avalanches per day, from 8th December 2019, to 30th May 2020, and from 8th December 2020, to 30th May 2021.

Number of avalanches	0	1	2	3	4	5	6	8
Frequency	294	29	17	4	4	2	1	1

Figure 11 presents the autocorrelation function (ACF) of daily avalanche count detected by the radar in Holmbuktura using data from both of the winter seasons. The correlation between days next to each other are significant, however it is not a strong correlation. At lags greater than 1 there is no significant correlations. Significance are indicated if the the ACF function crosses one of the blue lines.

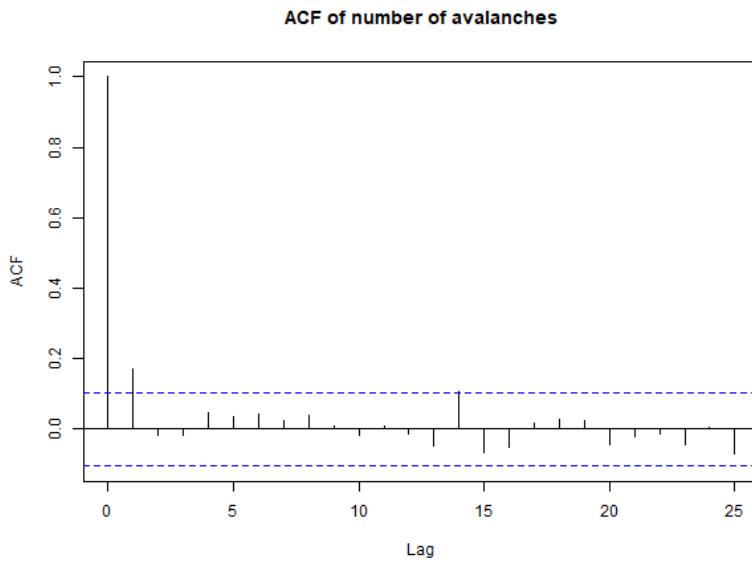


Figure 11: Plot showing the autocorrelation function of the radar detections of avalanches in Holmbuktura.

### 3.2.2 Avalanche data from satellites

The avalanche data collected from satellites are from Lavangsdalen. Lavangsdalen is selected due to the quality of the data, and because of the proximity to Holmbuktura, as shown in Figure 12.



Figure 12: Map of the position of Lavangsdalen relative to Holmbuktura. The straight-line distance between Holmbuktura and Lavangsdalen is 19 kilometers.

Source: Google Maps, 30th May 2024, <https://www.google.com/maps>

In the dataset, there are various information about the accuracy of the data, the estimated date, the avalanche release, and the surroundings of the avalanches as columns. One row in the dataset represents one unique observation of an avalanche release. In the analysis, avalanche releases, quality of the data as a letter from A to C, and the direction the mountain side of the avalanche releases are facing are used. Quality A means good quality, while Quality C means poor quality.

A frequency table of the number of avalanches per day can be found in Table 2. In total, there are 873 avalanches detected by satellites in Lavangsdalen. The maximum number of avalanches per day is 42. The maximum number of avalanches is shown in the time series plot of the data in Figure A.3 and A.4 in Appendix A. Further, the satellite time series plots also illustrates most of the detections happening in the start and in the end of the winter. This is due to wet snow avalanches being easier to observe (Eckerstorfer et al., 2017, p. 9), and wet snow avalanches occurring at high temperatures, which is more normal in the beginning and the end of the winter.

Table 2: Frequency table of satellite observations of avalanches per day, from 7th December 2019, to 30th May 2020, and from 7th December 2020, to 30th May 2021.

Number of avalanches	0	1	2	3	4	5-9	10-19	20+
Frequency	209	40	11	15	9	32	15	21

---

### 3.2.3 Weather forecast

The weather forecast is a lead time 1 forecast. It is called the mountain weather, and is originally from Varsom. It is valid for the whole date it is set, and it is published the day before. The forecast is published as quantiles with seven different quantiles for each weather variable. 0, 5, 25, 50, 75, 95, and 100 percent quantiles are included in the forecast. This means the values of the weather variable that 0, 5, 25, 50, 75, 95, and 100 percent of the region are expected to be below. In the analysis, the weather forecast for Lyngen are used. Weather parameters are chosen based on the risk factors discussed in Section 2.1. The following weather parameters are included in the analysis:

1. Quantiles of day-ahead forecast of daily precipitation. Measured in millimeters.
2. Quantiles of day-ahead forecast of daily average precipitation in the 400 squared kilometers with the most precipitation in the region. Measured in millimeters.
3. Quantiles of day-ahead forecast of daily average wind speed in the region. Measured in meters per second.
4. Quantiles of day-ahead forecast of daily snow depth. Measured in millimeters.
5. Quantiles of day-ahead forecast of daily temperature. Measured in Celsius.
6. Quantiles of day-ahead forecast of daily temperature. Measured in Celsius.
7. Quantiles of day-ahead forecast of daily average new snow in the region. Measured in millimeters.
8. Quantiles of day-ahead forecast of daily average new snow in the 400 squared kilometers with the most new snow in the region. Measured in millimeters.

The weather forecast for a date is separated into 5 different day-ahead forecasts for 5 different elevation zones. The elevation zones are intervals of 300 meters from 0 to 1500 meter above sea level. The forecasts for all elevation zones are used in the analysis. This is because what can be observed in Figure 10. The mountains in Holmbuktura reaches higher than 1000 meters above sea level, with peaks taller than 1500 meters above sea level.

The quantiles of day-ahead forecasts of precipitation and precipitation in the 400 squared kilometers with the most precipitation in Lyngen is identical for the different heights. The difference between the various quantiles of day-ahead precipitation are illustrated with an example week in Figure 13. An example week of precipitation in the area with the most precipitation are shown in Figure 14. It can be seen that the quantiles of precipitation in the area with the most precipitation are less spread out. The different quantiles are highly dependent.

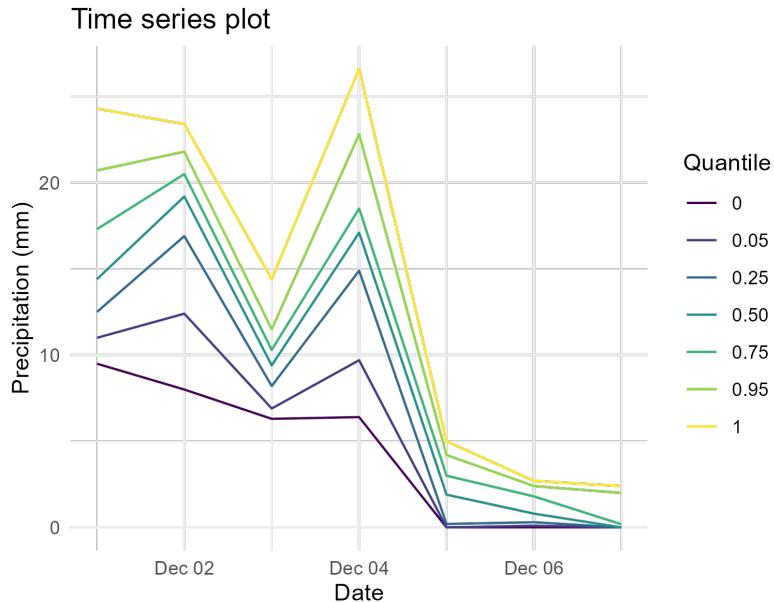


Figure 13: Time series plot of day-ahead precipitation forecast from 1st December 2019 to 7th December 2019.

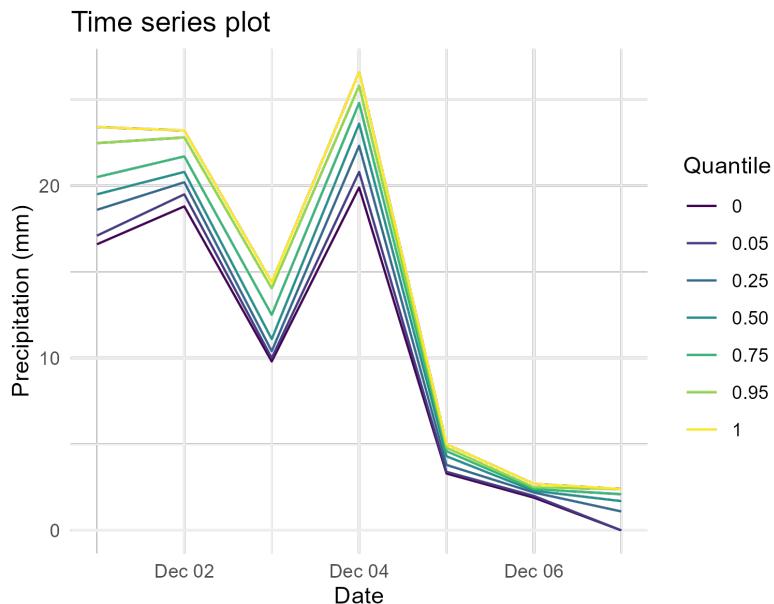


Figure 14: Time series plot of day-ahead forecast of precipitation for the area with the most precipitation from 1st December 2019 to 7th December 2019.

Preliminary explorative analysis revealed a suspicious forecast. Forecasted day-ahead quantiles of the precipitation and precipitation in the 400 squared kilometers with the most precipitation on 30th April 2020 are set to 0 in the analysis. This is because of all the forecasted quantiles being equal to 6280.5 millimeters, and this clearly wrong. The quantiles are set to 0, because the quantiles of new snow is identical to zero, while the quantiles of the temperature reaches below zero. This should not be the case if the quantiles of precipitation was different from zero.

---

The quantiles of temperature are illustrated in Figure 15. It can be seen that the different quantiles for the same height are highly dependent, and that this is also the case for the same quantiles for the two different heights. As expected, the temperature is lower at higher elevations.

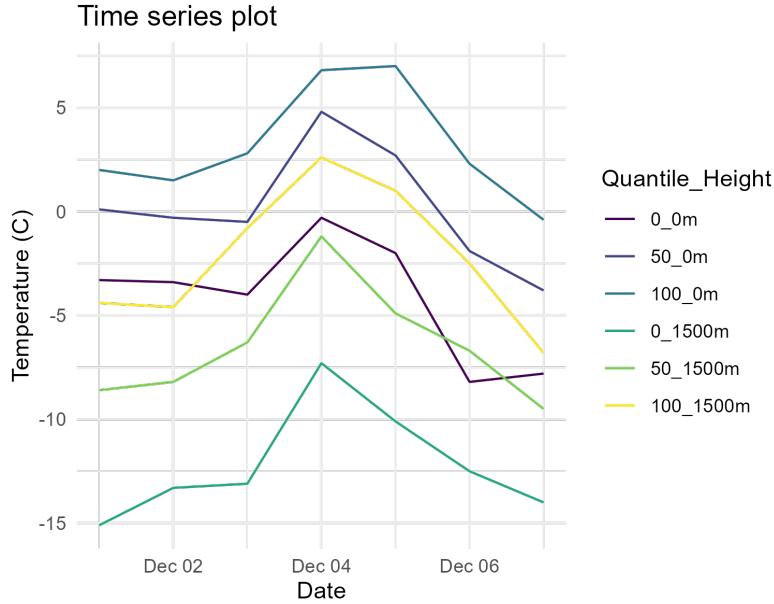


Figure 15: Time series plot of 0, 50 percent, and 100 percent quantiles of day-ahead temperature forecast for the elevation zones 0-300 and 1200-1500 meters above sea level. The time series is from 1st December 2019 to 7th December 2019.

The quantiles of wind for two different elevation zones are illustrated in Figure 16. The 0 quantile of wind is 0 m/s for the whole dataset and for all heights. The rest of the time series are highly dependent, but there are some irregularities.

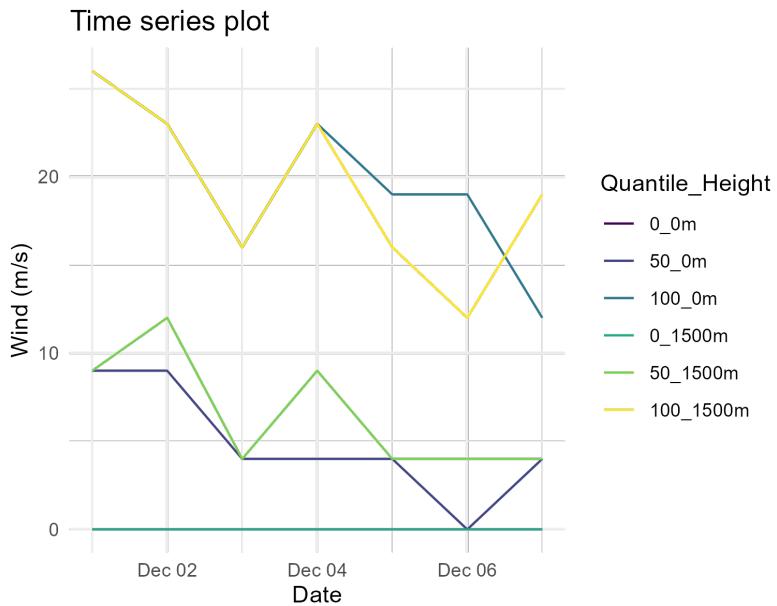


Figure 16: Time series plot of 0, 50 percent, and 100 percent quantiles of day-ahead wind forecast for the elevation zones 0-300 and 1200-1500 meters above sea level. The time series is from 1st December 2019 to 7th December 2019.

The quantiles of snow depth for two different elevation zones in the first winter are plotted in Figure 17. The dependence between the different quantiles are really high. There are also dependence between the different elevation zones, but this correlation is weaker.

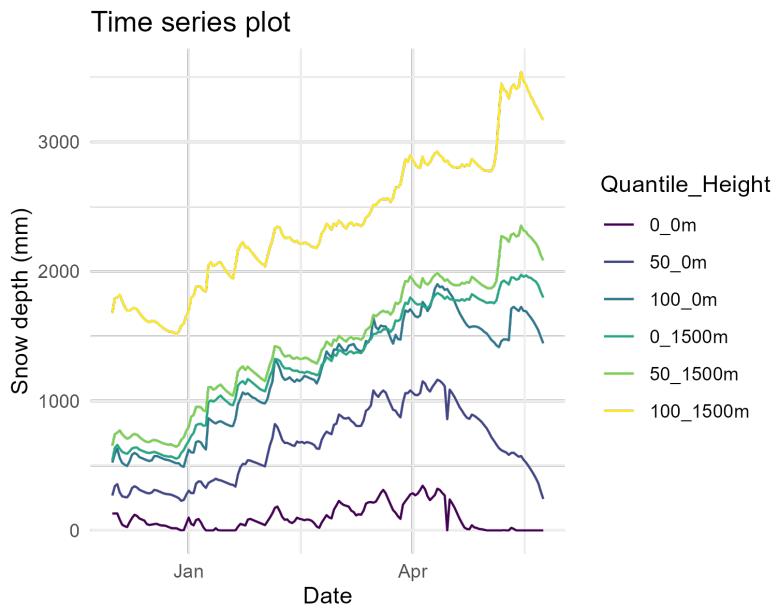


Figure 17: Time series plot of the quantiles of day-ahead snow depth forecast for 0-300 and 1200-1500 meters above sea level. The time series is from 1st December 2019 to 24th May 2020.

The same quantiles for different elevation zones are identical for day-ahead new snow forecasts. The first month of forecasted amount of new snow and forecasted amount of new snow in the area with the most forecasted new snow is plotted in Figure 18 and Figure 19. The conclusions from these time series plot are the same as the conclusions from the precipitation plots. It can be seen that the quantiles of new snow in the area with the most precipitation has quantiles with lower variance. The different quantiles are highly dependent.

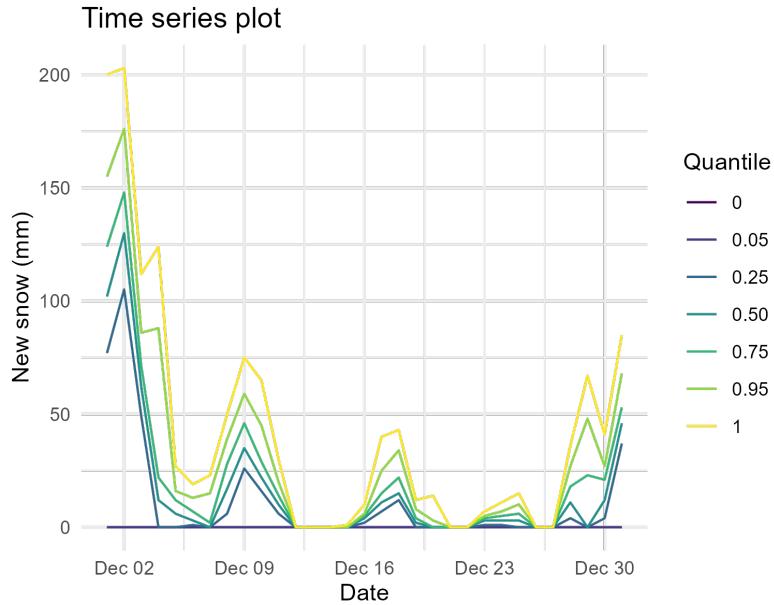


Figure 18: Time series plot of the quantiles of day-ahead new snow forecast. The time series is from 1st December 2019 to 31st December 2019.

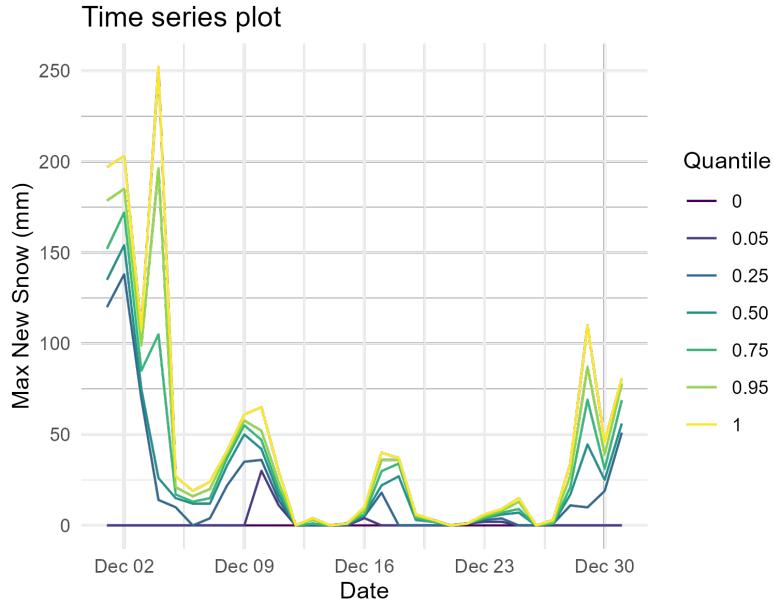


Figure 19: Time series plot of the quantiles of day-ahead forecast of new snow in the area with the most predicted new snow. The time series is from 1st December 2019 to 31st December 2019.

---

The data about the weather forecast is collected from a API from NVE through a Python library called Regobslib (Widforss, 2023).

### 3.2.4 Observed weather

Observed precipitation and snow depth at the weather station Ytre Holmbukt are used in the analysis. The position of Ytre Holmbukt can be seen in Figure 20. In this dataset, observation of snow depth from 9th May 2021 are missing. The missing value are set to 1 cm. This is done due to the snow depth being 2 cm the day before and 0 cm the day afterwards.



Figure 20: Map of the position of the weather station Ytre Holmbukt.

Source: Google Maps, 4th March 2024, <https://www.google.com/maps>

The weather observation data is collected from Seklima at the web page <https://seklima.met.no/>.

### 3.2.5 Danger level

In Norway today, avalanche warnings are issued as a manually set danger level on a scale from 1 to 5. The data about the danger level is, like the weather forecasts, collected from a API from NVE through a Python library called Regobslib (Widforss, 2023). The danger levels for Lyngen are used in the analysis.

---

### 3.2.6 Correlation analysis between types of data

It is important to explore the data and find relationships between the different variable types. These relations will be considered when variable selection are performed, and when inference are made.

In Figure A.5 in Appendix A it can be seen that there are strong correlations between some of the variable types. Another finding from Figure A.5 in Appendix A is the lack of evidence of considerable dependence between satellite data and other variables.

## 3.3 Choosing reduced set of explanatory variables

The reduced set of explanatory variables are chosen based on the theory from Section 2.1 about risk factors for avalanche formations, time series of the data in Section 3.2, the correlation plot in Figure A.5 in Appendix A, and the plots regarding the importance of the explanatory variables in Figure B.1 and B.2 in Appendix B. The explanatory variables chosen in the reduced set are:

1. Observed precipitation at the weather station Ytre Holmbukt with lag 1 (Precip\_1\_). It can be seen in Figure B.1 and B.2 in Appendix B that a lot of the precipitation related variables with lag 1 have a high variable importance. According to Figure B.2 in Appendix B, the observed precipitation at the weather station Ytre Holmbukt with lag 1 is one of the most important variables.
2. The 0-quantile of day-ahead forecast of precipitation in the area with the most precipitation (precip\_max\_150). This is the most important variable according to both B.1 and B.2 in Appendix B.
3. The 1-quantile of day-ahead new snow forecast (new\_snow100\_150). This is one of the most important variables according to B.1 in Appendix B. This variable also chosen to get an precipitation related explanatory variable as different as possible from the 0-quantile of day-ahead forecast of precipitation in the area with the most precipitation. The time series of the data in Section 3.2 and the correlation plot in Figure A.5 in Appendix A suggests strong correlation between precipitation related variables at the same time lag.
4. The 0.25-quantile of the day-ahead snow depth forecast in between 600 and 900 meters above sea level (snow\_depth25\_750). Various variables related to the snow depth are picked as important by Figure B.1 and B.2 in Appendix B. As shown in Figure 17, there are strong correlation between snow depth related variables, and also at different time lags, so it less important which quantiles that are chosen. The reason for choosing the forecast between 600 and 900 meters above sea level is that avalanches are typically released far up in the mountain side.

- 
5. The 0.95-quantile of day-ahead temperature forecast between 600 and 900 meters above sea level with lag 1 (temp\_95\_1). Various variables related to the temperature at different lags are picked as important by Figure B.1 and B.2 in Appendix B. From the same figures we can observe a trend of high quantiles and elevation zones being chosen by the variable importance plots.
  6. The 1-quantile of day-ahead wind forecast between 600 and 900 meters above sea level (wind100\_750). This variable is also chosen based on a trend in Figure B.2 in Appendix B that various high quantiles are listed as important variables. The same altitude is used as the rest of the variables.

---

## 4 Modelling

This thesis models the observed avalanches as a univariate response of daily avalanche count. The explanatory variables are variables that are available the day before. The goal is to predict the avalanche count given information from the day before, and to do statistical inference. There are 352 unique observations of daily avalanche count used for model fitting. The observations are from 8th December to 31st May in the two winter seasons. 1st December to 7th December is excluded to allow lag 1 to 7 to be explanatory variables.

Four different types of models are fitted. The models used to predict the response are:

1. Tree-based models.
2. Models fitted using the method of nearest neighbours.
3. Generalized linear models.
4. Ensemble models.

All of the models are evaluated by being compared to three baseline models. These models are generalized linear models fitted with no explanatory variables, danger level as response, and observed avalanche count with lag 1.

### 4.1 Sets of explanatory variables

The models were fitted using two different selections of explanatory variables. In one of the sets, the full set of explanatory variables are included. The other set consists of a reduced set of these variables.

The full set of explanatory variables includes all relevant information in the data available the day before. In total there are 1107 explanatory variables. Information from these sources the last week before the day of interest are included in the analysis. The available data are introduced in Section 3.2, and includes the following information:

1. Quantiles of the weather forecasts for different elevation zones.
2. Weather observations from Ytre Holmbukt.
3. Number of avalanche release detected by satellites.
4. Number of radar detections from Holmbuktura.

---

This means that the covariate matrix of the full set of explanatory variables,  $\mathbf{X}_{\text{full}}$ , would be a  $352 \times 1107$  matrix with rows as observations and columns as the different covariate vectors containing information about the explanatory variables.  $\mathbf{X}_{\text{full}}$  is illustrated in (17).

$$\mathbf{X}_{\text{full}} = \begin{bmatrix} x_{11} & x_{12} & \cdots \\ x_{21} & x_{22} & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \quad (17)$$

Each column

$$\mathbf{x}_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \end{bmatrix} \quad (18)$$

represents the the observations of the  $j$ -th explanatory variable, for  $j=1, 2, \dots, 1107$ .

The reduced set of explanatory variables are selected based on the findings from Section 3.3. The variables selected are:

1. Observed precipitation at the weather station Ytre Holmbukt with lag 1 (Precip\_1\_-).
2. The 0-quantile of day-ahead forecast of precipitation in the area with the most precipitation (precip\_max\_150).
3. The 1-quantile of day-ahead new snow forecast (new\_snow100\_150).
4. The 0.25-quantile of the day-ahead snow depth forecast in between 600 and 900 meters above sea level (snow\_depth25\_750).
5. The 0.95-quantile of day-ahead temperature forecast between 600 and 900 meters above sea level with lag 1 (temp\_95\_1).
6. The 1-quantile of day-ahead wind forecast between 600 and 900 meters above sea level (wind100\_750).

$\mathbf{X}_{\text{reduced}}$  is a  $352 \times 6$  matrix as seen in (19).

$$\mathbf{X}_{\text{reduced}} = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} & x_{16} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} & x_{26} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \quad (19)$$

Each column

$$\mathbf{x}_j = \begin{bmatrix} x_{1j} \\ x_{2j} \\ \vdots \end{bmatrix} \quad (20)$$

represents the  $j$ -th explanatory variable, for  $j = 1, 2, \dots, 6$ .

---

## 4.2 Tree models

### 4.2.1 The models

There are in total fitted eight tree-based models. Four of them are fitted using the smaller set of explanatory variables, while the other four are fitted using all the explanatory variables. For each set of explanatory variables, two of the models are ordinary decision trees, while the two others has a response which are calculated from growing multiple trees.

The simple decision trees has sum of squared errors (SSE) as splitting criterion. This approach involves selecting splits minimizing the SSE at each step, thereby refining the tree's accuracy. The algorithm evaluates all possible splits for each feature, choosing the one that offers the most significant reduction in SSE. These trees are fitted in R using the library called rpart (Atkinson, n.d.).

For each set of explanatory variables, there is one tree fitted using the standard algorithm and one tree fitted which are a pruned version of the original tree. The performance of different number of terminal nodes are evaluated using 5-fold cross-validation of trees with different number of terminal nodes, and then the pruned tree are chosen from the number of terminal nodes that gives the lowest RMSE. 5-fold cross-validation divides the data into five folds. Four of the folds are used for training the model, while the last one is used for evaluation. Pruning is performed using the library (Kuhn, 2024).

The random forest model are fitted using the library randomForest (Breiman and Cutler, n.d.) in R. First, the numbers of explanatory variables considered at each split in the random forest are estimated by 5-fold cross-validation using the caret library (Kuhn, 2024). After this, the trees are grown and averaged. The random forest models grows 500 trees. This is to make sure that each day in the dataset are part of the bootstrapped samples, which the trees are fitted from, multiple times.

The last tree-based method used are boosting. The boosted models are fitted using the library gbm, which is short for gradient boosted models (Ridgeway, n.d.). The parameters of the boosted models are also estimated by 5-fold cross-validation using the caret library (Kuhn, 2024). The parameters that are estimated using cross-validation are the interaction depth, meaning the maximum depth of interactions between variables allowed, and the number of trees. The shrinkage parameter is held constant at 0.1.

### 4.2.2 Evaluating the models

The pruned tree, the random forest and the boosting model are all evaluated by looking at the CRPS. Tree-based models give a point prediction, and the CRPS will therefore be the area between two indicator functions. An example on how the indicator functions can look are illustrated in Figure 21. Model evaluation using CRPS is performed by using the R-library scoringRules (Krüger, 2023).

---

The same three trees are also evaluated by performing 5-fold cross-validation and looking at root mean cross validation error. The library caret (Kuhn, 2024) are used for the cross-validation of tree models.

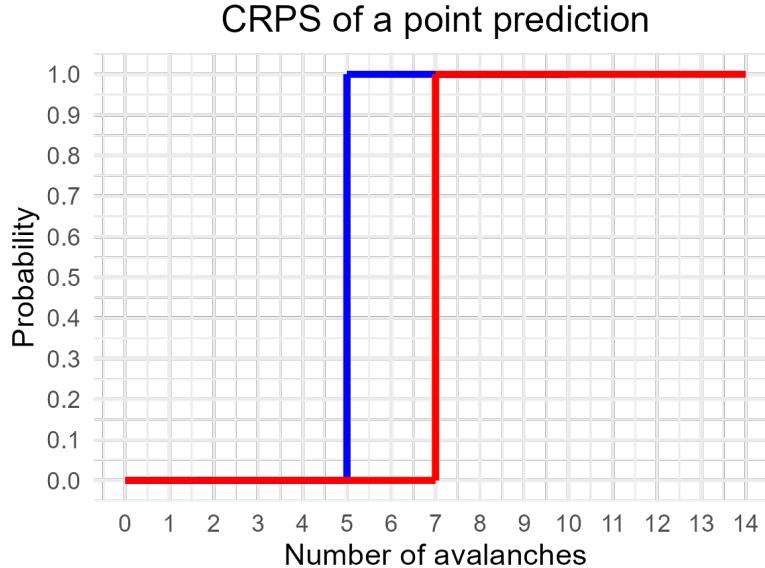


Figure 21: Illustration of CRPS of point predictions. The blue graph shows the CDF of a point prediction equal to 5, while the red graph shows the the indicator function  $\mathbb{I}\{y \geq 7\}$ .

In addition, the variable importance of the different explanatory variables in the random forests and the boosted models are evaluated. This is done for random forests by looking at increase in node purity, measured in decrease in RSS, using the library (Breiman and Cutler, n.d.). For the boosting models, this is done by looking at the relevant information gained from each variable using the library (Ridgeway, n.d.). These evaluation techniques are important when selecting the smaller set of explanatory variables.

## 4.3 Models fitted using method of nearest neighbours

### 4.3.1 The models

The models fitted using the method of nearest neighbours uses the 2, 5 and best number of nearest neighbours to predict the avalanche count. The best number of nearest neighbours are calculated by performing 5-fold cross-validation. The models are fitted using  $\mathbf{X}_{\text{full}}$  and  $\mathbf{X}_{\text{reduced}}$ . The formula for predicting the response using the method of nearest neighbours are shown in (13) If the 2 nearest neighbours are used, and if  $x_0$  is the vector of the explanatory variables of interest and  $N_i$  the i-th neighbour among the nearest neighbours, then the response of interest,  $f(x_0)$ , is

---

calculated by following (21).

$$f(x_0) = \frac{1}{2} \sum_{i=1}^2 N_i \quad (21)$$

### 4.3.2 Evaluating the models

The evaluation of models fitted using the method of nearest neighbours are performed using CRPS and RMSE. The response is, as with trees, assumed to be a point prediction. Therefore, the CRPS will be the area between two indicator functions. In Figure 21, an example of the functions used in the calculation of CRPS is illustrated. CRPS is calculated using the library (Krüger, 2023). RMSE is calculated using 5-fold cross-validation and the library (Kuhn, 2024), which is also the case for the RMSE of the tree-based models.

## 4.4 Generalized linear models

### 4.4.1 The models

Three generalized linear models are fitted using  $\mathbf{X}_{\text{reduced}}$  and log link. The difference between the models are the systematic component. All of the models are assumed to have a Poisson distributed response.

The first generalized linear model is a simple model where each of the explanatory variables are represented once in the systematic component. This model has the systematic component shown in equation (22).

$$\eta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_6 x_{i6}, \quad (22)$$

where  $\beta_1, \beta_2, \dots$ , and  $\beta_6$  are the regression parameters and  $x_{i1}, x_{i2}, \dots$ , and  $x_{i6}$  are the explanatory variables.

The second generalized linear model also has interaction terms between all the explanatory variables of depth two. The systematic component are shown in equation (23).

$$\eta_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_6 x_{i6} + \beta_{12} x_{i1} x_{i2} + \beta_{13} x_{i1} x_{i3} + \cdots + \beta_{56} x_{i5} x_{i6} \quad (23)$$

The last model contains polynomials of power three for every explanatory variable. This systematic component is shown in (24).

$$\eta_i = f_{i1}(x_{i1}) + f_{i2}(x_{i2}) + \cdots + f_{i6}(x_{i6}), \quad (24)$$

---

where  $f_{i1}(x_{i1}), f_{i2}(x_{i2}), \dots, f_{i6}(x_{i6})$  are the polynomials expressed as functions. For explanatory variable  $j$ , the polynomial  $f_{ij}(x_{ij})$  is shown in (25).

$$f_{ij}(x_{ij}) = \beta_{1j}x_{ij} + \beta_{2j}x_{ij}^2 + \beta_{3j}x_{ij}^3 \quad (25)$$

The generalized linear models are fitted using functions from base R.

#### 4.4.2 Evaluating the models

As is the case for both tree-based models and nearest neighbours models, the generalized linear models are evaluated by calculating the CRPS and RMSE. The difference is that the generalized linear models gives a probabilistic forecast. The library scoringRules (Krüger, 2023) are still used for calculations, but the general linear models do not issue a point prediction. The predictions of the generalized linear models is a probabilistic forecast which are assumed to be Poisson distributed. This means that the CRPS, expressed as the area between the CDF of the prediction and an indicator function  $\mathbb{I}\{y \geq y_{obs}\}$ , would look like the area between the functions in Figure 5. The RMSE is calculated using (Kuhn, 2024) like for the other models.

Further, the generalized linear models are evaluated using histograms of the probability integral transforms, called PIT-diagrams, discussed in Section 2.2.2. The PIT-diagrams test if the assumption of Poisson distributed responses holds. The PIT-diagrams are created using the R-library tscount (Liboschik, 2023). Tscount uses a non-randomized uniform version of the pit-diagram, described by a Czado et al. (2009, p. 2), to account for the discrete outcomes.

### 4.5 Ensemble model

#### 4.5.1 The model

The ensemble model produces a probabilistic forecast. The response is assumed to be Poisson distributed. Model fitting are performed using the avalanche forecast from the random forest model, the model fitted using the 20 nearest neighbours, and the simple generalized linear model as explanatory variables. All three of the models are fitted using  $\mathbf{X}_{\text{reduced}}$  and forecasted on training data. The systematic component of the ensemble model can be seen in

$$\eta_i = \beta_{RF}X_{RF,i} + \beta_{GLM}X_{GLM,i} + \beta_{20NN}X_{20NN,i} + \dots, \quad (26)$$

where  $X_{RF}$  are the forecast by the random forest,  $X_{GLM}$  are the forecast by the generalized linear mode,, and  $X_{20NN}$  are the forecast from the model fitted using the method of nearest neighbours. An identity link function is used.

---

#### 4.5.2 Evaluating the model

The ensemble model is evaluated in the same way as the generalized linear models. The evaluation of the generalized linear models is discussed in Section 4.4.2.

### 4.6 Evaluating models on example dates

The resulting models are illustrated by comparing the results of three different days with different avalanche count detections. The days are called day 1, day 2 and day 3. The cumulative distribution function of the models chosen in the ensemble model in Section 5.4.1.

1. Day 1 is 16th December 2019. On this date, there were no avalanches detected by the radar in Holmbuktura.
2. Day 2 is 13th January 2020, where there was 1 avalanche detected in Holmbuktura. This avalanche also hit the road in Holmbuktura.
3. Day 3 has 3 radar observations of avalanches in Holmbuktura, where one of the avalanches hit the road. The date are 4th December 2019.

The values of the explanatory variables in  $\mathbf{X}_{\text{reduced}}$  are shown in Table 3.

Table 3: The values of the explanatory variables at the example dates.

	precip_max0_150	new_snow100_150	temp95_1_750
Day 1	1.90	30.00	-6.90
Day 2	7.80	70.00	-0.80
Day 3	9.20	140.00	-0.10

	wind100_750	Precip_1	snow_depth25_750
Day 1	26.00	4.80	513.00
Day 2	16.00	5.40	895.67
Day 3	26.00	5.00	555.00

---

## 5 Results

In this section, the results for each of the model classes, i.e. tree models, models fitted using the method of nearest neighbors, and generalized linear models are presented. For each model class the two sets of explanatory variables,  $\mathbf{X}_{\text{full}}$  and  $\mathbf{X}_{\text{reduced}}$ , are used. Further, the results for the different model classes are compared by looking at avalanche count predictions in Section 5.6.1 on three different days with different number of avalanche count detections, and by comparing the models with predictions from baseline models in Section 5.6.2.

### 5.1 Tree models

#### 5.1.1 Simple regression trees

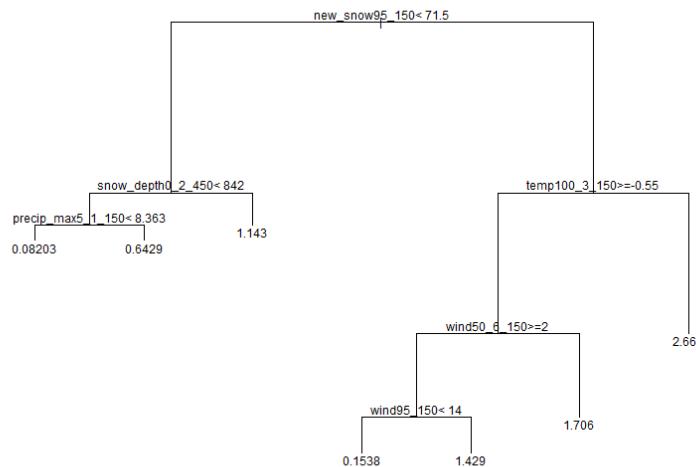


Figure 22: Regression tree fitted using  $\mathbf{X}_{\text{full}}$ . Output values are predicted avalanche count.

A simple regression tree fitted to the data using all explanatory variables is illustrated in Figure 22. The first split is determined by the 0.95-quantile of the predicted amount of new snow (new\_snow95\_150). Other variables used in the analysis are:

1. The 0-quantile of day-ahead snow depth forecast between 300 and 600 meters above sea level with lag 2 (snow\_depth0\_2\_450).
2. The 0.05-quantile of day-ahead precipitation forecast with lag 1 in the area with the most predicted precipitation (precip\_max5\_1\_150).
3. The 1-quantile of day-ahead temperature forecast between 0 and 300 meters below sea with lag 3 (temp100\_3\_150).

- 
- 4. The 0.5-quantile of day-ahead wind forecast between 0 and 300 meters above sea with lag 6 (wind5\_6\_150).
  - 5. The 0.95-quantile of day-ahead wind forecast between 0 and 300 meters above sea level (wind95\_150).

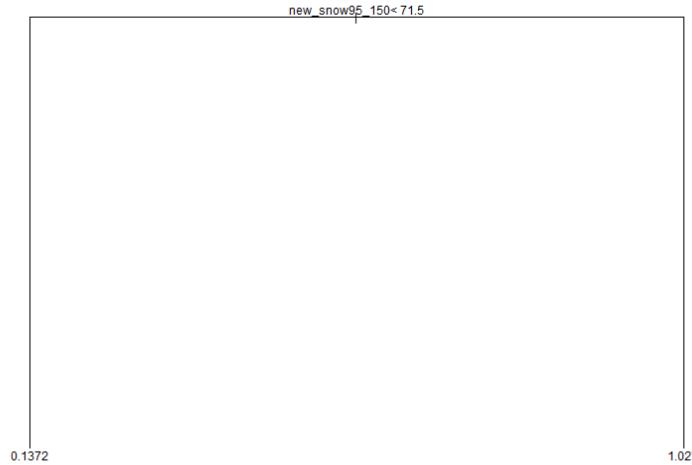


Figure 23: Pruned regression tree fitted using  $\mathbf{X}_{\text{full}}$ . Output values are predicted avalanche count.

Using cross-validation, the simple regression tree in Figure 22 is pruned using the method presented in Section 4.2.1. The pruned tree is illustrated in Figure 23. It can be seen that the pruned tree only has one split, meaning the 0.95-quantile of the predicted amount of new snow is the only variable in use. The pruned tree suggests that if the 0.95-quantile of the predicted amount of new snow is less than 71.5mm, then an avalanche count of 0.1372 is predicted. Else, an avalanche count of 1.027 is predicted.

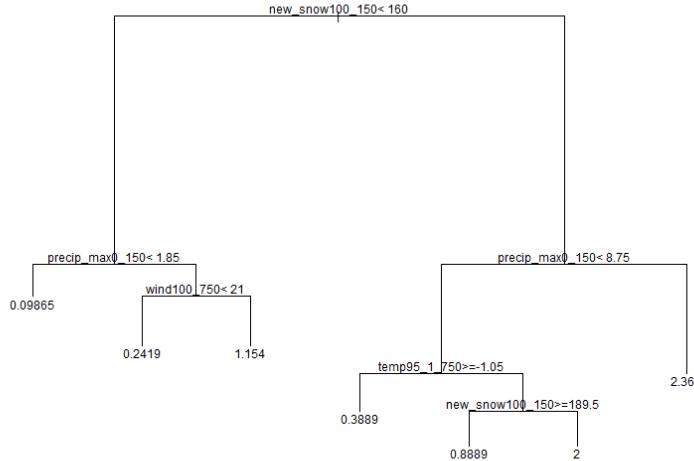


Figure 24: Regression tree fitted using  $\mathbf{X}_{\text{reduced}}$ . Output values are predicted avalanche count.

A simple regression tree fitted to the data using the reduced set of explanatory variables is illustrated in Figure 22. Four out of six explanatory variables are in use in the simple regression tree. The first split is determined by the 1-quantile of the predicted amount of new snow (`new_snow100_150`). Other variables used in the analysis are:

1. The 0-quantile of day-ahead precipitation forecast in the area with the most predicted precipitation (`precip_max0_1_150`).
2. The 0.95-quantile of day-ahead temperature forecast between 600 and 900 meters below sea level with lag 1 (`temp95_1_750`).
3. The 1-quantile of day-ahead wind forecast between 600 and 900 meters above sea level (`wind100_750`).

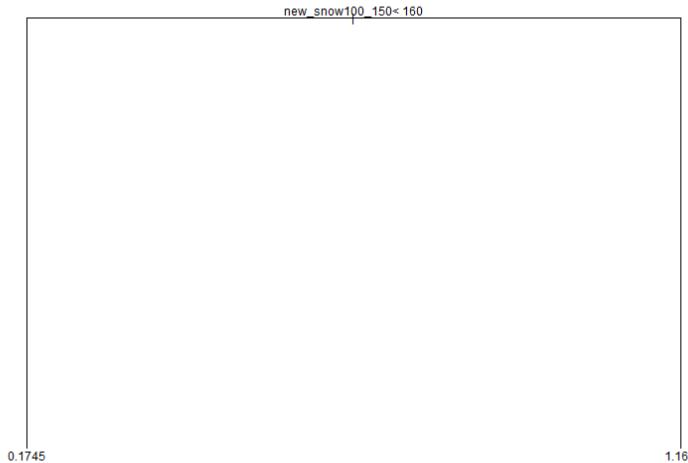


Figure 25: Pruned regression tree fitted using  $\mathbf{X}_{\text{full}}$ .

Using cross-validation, the simple regression tree in Figure 22 is pruned. The pruned tree is illustrated in Figure 25. It can be seen that the pruned tree only has one split, meaning that the 1-quantile of the predicted amount of new snow is the only variable in use. The pruned tree suggests that if the 1-quantile of the predicted amount of new snow is less than 160mm, then an avalanche count of 0.1745 is predicted. Else, an avalanche count of 1.167 is predicted.

### 5.1.2 Random forests and boosting models

Two random forests and two boosting models were fitted. One random forest and one boosting model were fitted using  $\mathbf{X}_{\text{full}}$ , while the other random forest and boosting model were fitted using  $\mathbf{X}_{\text{reduced}}$ .

The increase in node purity of the variables in the random forest fitted using  $\mathbf{X}_{\text{full}}$  indicates predicted amount of new snow and precipitation and the predicted amount of new snow and precipitation in the area with the most precipitation being the most important variables. Snow depth, temperature and predicted amount of the precipitation variables and new snow variables the day before are also important variables. In Figure B.1 in Appendix B, the importance of the 40 most important variables in the random forest fit using  $\mathbf{X}_{\text{full}}$  are plotted.

The most important variables of the boosting model fitted using  $\mathbf{X}_{\text{full}}$  are, as for random forest, predicted amount of the precipitation variables and new snow variables. In addition, observed precipitation and predicted amount of precipitation variables the day before also have a high importance. The boosting fit involves a greater variety of important variables than the random forest fit. In Figure B.2 in Appendix B, the importance of the 40 most important variables in the boosting fit using  $\mathbf{X}_{\text{full}}$  are plotted.

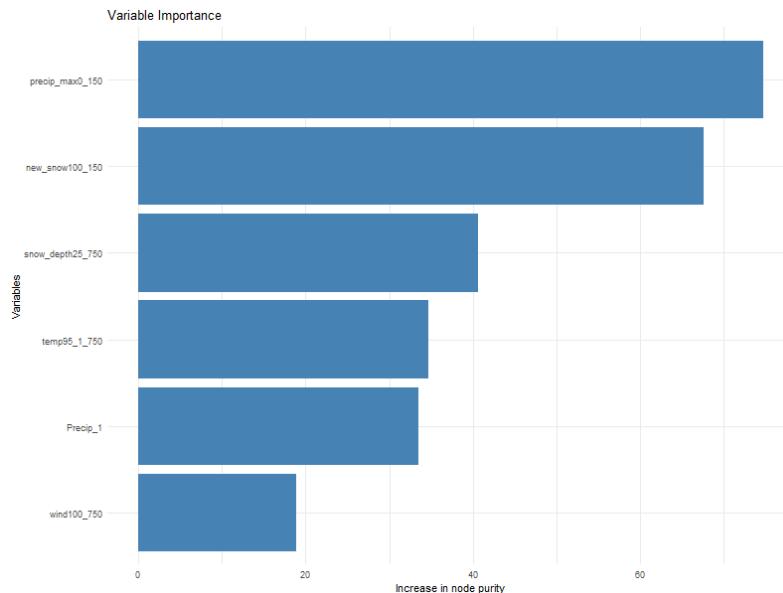


Figure 26: Variable importance plot of the random forest model fitted using  $\mathbf{X}_{\text{reduced}}$

In Figure 26, the importance of the explanatory variables in the random forest fit using  $\mathbf{X}_{\text{reduced}}$  are plotted. The plot shows that the precipitation and new snow forecasts are the two most important variables, while the wind variable seems less important.

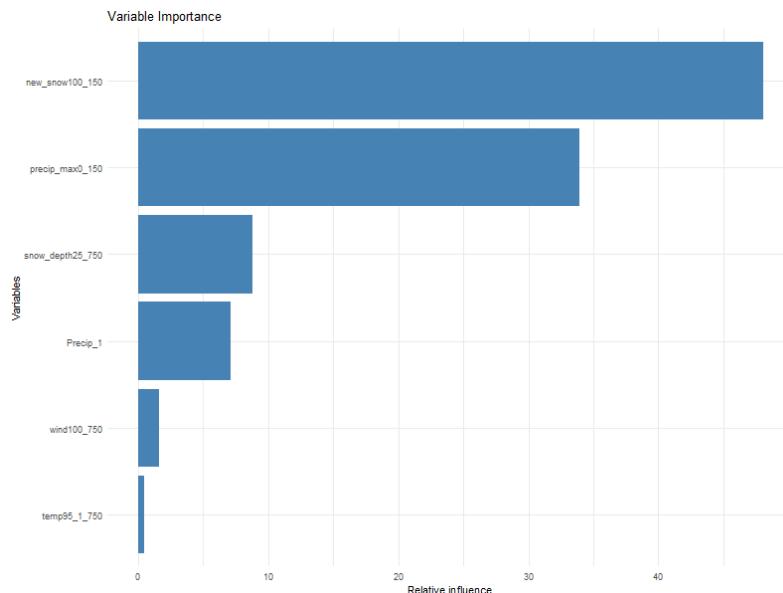


Figure 27: Variable importance plot of the boosting model fitted using  $\mathbf{X}_{\text{reduced}}$ .

In Figure 27, the importance of the explanatory variables in the boosting fit using  $\mathbf{X}_{\text{reduced}}$  are plotted. This plot also shows that he precipitation and new snow forecast are the two most important variables, while the wind variable seems less important. The difference is that the relative influence of the temperature variable are also very low.

### 5.1.3 Performance of tree-based models

The CRPS and RMSE of all the tree-based models are calculated to evaluate the performance of the models.

Table 4: Performance of the pruned tree, random forest and boosting model fitted using both  $\mathbf{X}_{\text{full}}$  and  $\mathbf{X}_{\text{reduced}}$ .

	Pruned $X_{\text{full}}$	RF $X_{\text{full}}$	Boost $X_{\text{full}}$	Pruned $X_{\text{reduced}}$	RF $X_{\text{reduced}}$	Boost $X_{\text{reduced}}$
CRPS	0.45	0.26	0.31	0.46	<b>0.23</b>	0.42
RMSE	0.91	0.89	0.89	0.89	<b>0.84</b>	0.91

The CRPS, defined in (5), and the RMSE, defined in (6), of the pruned trees, the boosting models and the random forests fitted using  $\mathbf{X}_{\text{full}}$  and  $\mathbf{X}_{\text{reduced}}$  are presented in Table 4. The CRPS indicates the random forests models performing the best and the predictions from the pruned tree model being less accurate than the prediction from the boosted model and random forest model. The RMSE gives the same conclusion. The RMSE also suggests the random forest model performing better using  $\mathbf{X}_{\text{reduced}}$ , and this is also supported by the findings from the CRPS. The CRPS of the boosting model increases when using  $\mathbf{X}_{\text{reduced}}$ , but the RMSE says the opposite. The CRPS and the RMSE also prefers different pruned tree models.

## 5.2 Models fitted using the method of nearest neighbours

The best number of neighbours minimizing the RMSE was found to be 9 for the full dataset and 20 for the reduced dataset using 5-fold cross-validation.

Table 5: Performance of models fitted using the 2, 5, 9, and 20 nearest neighbours. Both  $\mathbf{X}_{\text{full}}$  and  $\mathbf{X}_{\text{reduced}}$  are used for fitting the models with 2 and 5 nearest neighbours. However, only  $\mathbf{X}_{\text{full}}$  is used for fitting the model with 9 nearest neighbours, and only  $\mathbf{X}_{\text{reduced}}$  is used for fitting the model with 20 nearest neighbours.

	2NN $X_{\text{full}}$	5NN $X_{\text{full}}$	9NN $X_{\text{full}}$	2NN $X_{\text{reduced}}$	5NN $X_{\text{reduced}}$	20NN $X_{\text{reduced}}$
CRPS	0.26	0.39	0.45	<b>0.24</b>	0.38	0.42
RMSE	1.06	0.90	<b>0.86</b>	1.00	0.93	0.88

In Table 5, the CRPS and the the RMSE of the models fitted using the method of nearest neighbours are shown. The RMSE are naturally lowest for the best models, which are picked based on the lowest RMSE. The RMSE also suggests that using the 5 nearest neighbours are better than using the 2 nearest best neighbours, both

---

when using  $\mathbf{X}_{\text{full}}$  and  $\mathbf{X}_{\text{reduced}}$ . The CRPS increases with the number of neighbours, meaning the CRPS indicates the model using the 2 nearest neighbours being the best model.

### 5.3 Generalized linear models

The generalized linear models are fitted only using the reduced set of explanatory variables.

Table 6: P-values and estimates of the parameters of the simple generalized linear model.

	P-values	Coefficients
(Intercept)	$3.60 \times 10^{-9}$	-2.28
precip_max0_150	$1.74 \times 10^{-7}$	0.1
new_snow100_150	$1.46 \times 10^{-5}$	$2.97 \times 10^{-3}$
temp95_1_750	0.11	-0.05
wind100_750	0.93	$1.37 \times 10^{-3}$
Precip_1	0.21	0.02
snow_depth25_750	0.26	$2.44 \times 10^{-4}$

The p-values and the parameter estimates of the simple generalized model can be found in Table 6. Table 6 indicates that the only parameters, besides the intercept, being significant are the positive coefficients of precip\_max0\_150 and new\_snow100\_150 using 95 percent confidence level. It is statistically significant that avalanche count increases with the 1-quantile of day-ahead new snow forecast and the 0-quantile of day-ahead forecast of precipitation in the area with the most precipitation.

In the generalized linear model with interaction terms precip\_max0\_150 is the only explanatory variable of interaction depth one being significant using 95 percent significance level. However, there are also a few of the explanatory variables of the interaction terms that are significant, but most of the terms are not significant. All the p-values of the model fitted with interaction terms can be found in in Table 15 in Appendix D.

In the GLM with polynomials all the explanatory variables except Precip\_1 and wind100\_750 have significant parameters using 95 percent significance level. This is shown by the p-values in Table 16 in Appendix D.

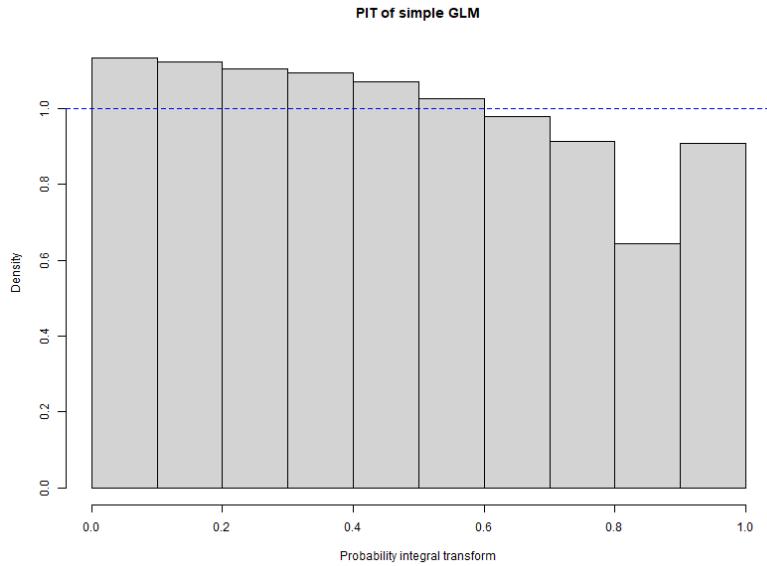


Figure 28: PIT-diagram of simple generalized linear model. All training data is used.

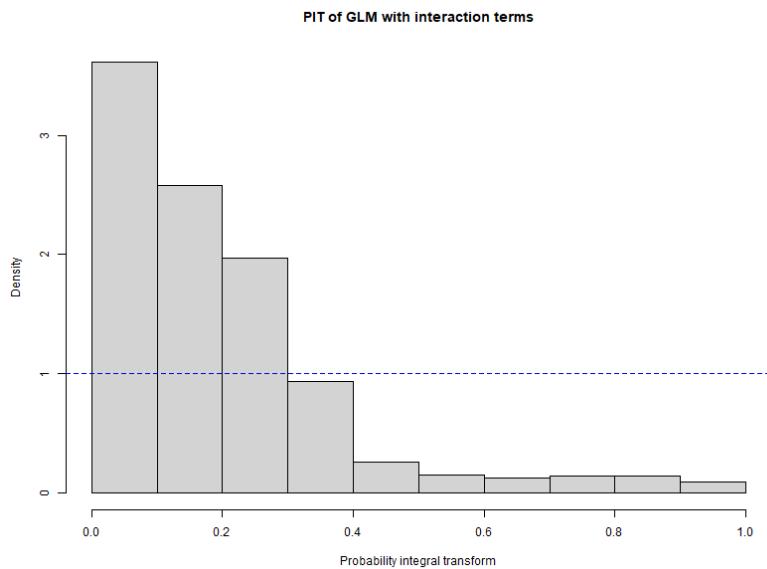


Figure 29: PIT-diagram of generalized linear model with interactions terms.

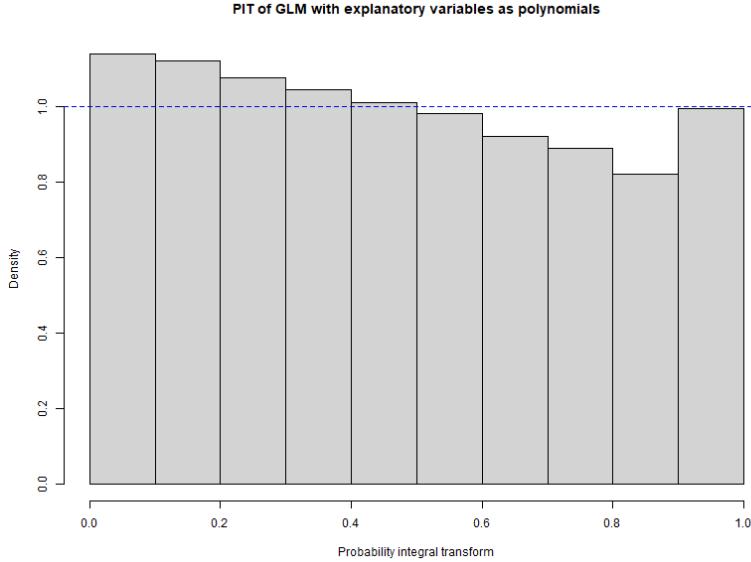


Figure 30: PIT-diagram of generalized linear model with polynomials.

The PIT-diagrams should be uniform if the models are correctly calibrated. The PIT-diagrams in Figure 28, 29, and 30 suggests all three generalized linear models tending to observe a lower avalanche count than what would be expected if the model was correctly calibrated. According to Figure 28 and 30, the simple generalized linear model and the generalized linear model are to some extent calibrated, the diagrams are not too far from uniform. The diagram in Figure 28 and 30 are very similar. The diagram in Figure 29 are far from uniform. Given the model, there are observed way fewer avalanches than what would be expected if the model was well calibrated.

Table 7: Performance of the various generalized linear models.

	Simple	With interaction terms	With polynomials
CRPS	0.27	0.82	<b>0.26</b>
RMSE	<b>0.93</b>	1.84	$4.94 \times 10^4$

By looking at the CRPS of the generalized linear models in Table 7, it can be seen that the simple generalized linear model and the generalized linear model with polynomials are favourable. The model with interaction terms performs far worse, like seen in the PIT-diagram in Figure 29. However, the RMSE of the generalized linear models indicates the simple generalized linear model being the best model, while the model with polynomial terms fails horribly.

---

## 5.4 Ensemble model

### 5.4.1 Choosing ensemble model

The models being the explanatory variables in the ensemble model are selected based on the CRPS and RMSE. The models chosen are:

1. The random forest fitted using  $\mathbf{X}_{\text{reduced}}$ . This model has the lowest CRPS and RMSE among the trees.
2. The model fitted using the 9 nearest neighbours and  $\mathbf{X}_{\text{full}}$ . This model has the lowest RMSE. The CRPS is less important for nearest neighbours models, since the CRPS increases with the number of neighbours.
3. The simple generalized linear model. The RMSE is the lowest, and the CRPS is just slightly higher than the model fitted using polynomials.

### 5.4.2 Ensemble model fit

The ensemble model are then fitted using data from the forecasts of the random forest model, the model fitted using the 20 nearest neighbours, and the simple generalized linear model as explanatory variable on the training data used for the same models.

Table 8: P-values and estimates of the parameters in the ensemble model.

	P-values	Coefficients
(Intercept)	0.21	0.02
glm_pred	0.17	-0.15
rf_pred	$6.80 \times 10^{-21}$	1.04
knn_pred	0.93	$-2.16 \times 10^{-3}$

From the p-values and estimates in Figure 8, it can be seen that the random forest model are chosen as the most important model for predicting the response. The random forest prediction are the only significant explanatory variable. The prediction from the generalized linear model is negative, while the prediction from the model fitted using the method of nearest neighbours is nearly equal to zero. This indicates that a model with only the prediction of the random forest as explanatory variable could perform better on new data, and be easier to interpret.

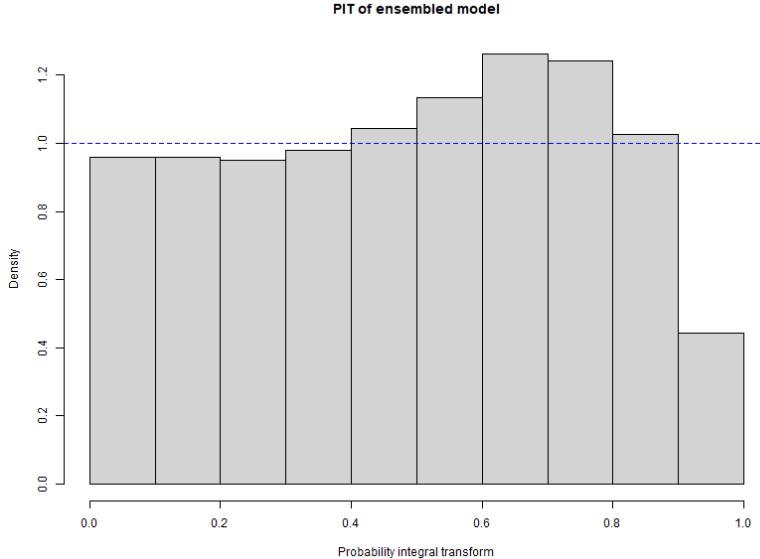


Figure 31: PIT-diagram of ensemble model

The PIT-diagram for the ensemble model suggests that the model is generally well-calibrated. However, there are noticeable deviations. Specifically, observations of the most extreme avalanche counts are fewer than expected, while there are more observations of high but not maximal avalanche counts than would be anticipated if the model were perfectly calibrated.

Table 9: Performance of the ensemble model

Ensemble model	
CRPS	0.11
RMSE	0.42

Table 9 shows the CRPS and the RMSE of the ensemble model.

## 5.5 Ensemble model with random forest prediction as the only explanatory variable

An ensemble model with random forest prediction as the only explanatory variable are also fitted. This is because the random forest predictions were the only significant explanatory variable of the ensemble model presented in Section 5.4.

---

Table 10: P-values and estimates of the parameters of the ensemble model with random forest as the only explanatory variable.

	P-values	Coefficients
(Intercept)	$1.71 \times 10^{-27}$	$-1.15 \times 10^{-3}$
rf_pred	$1.71 \times 10^{-27}$	1

From the p-values and estimates in Figure 10, it can be seen that the random forest variable is significant

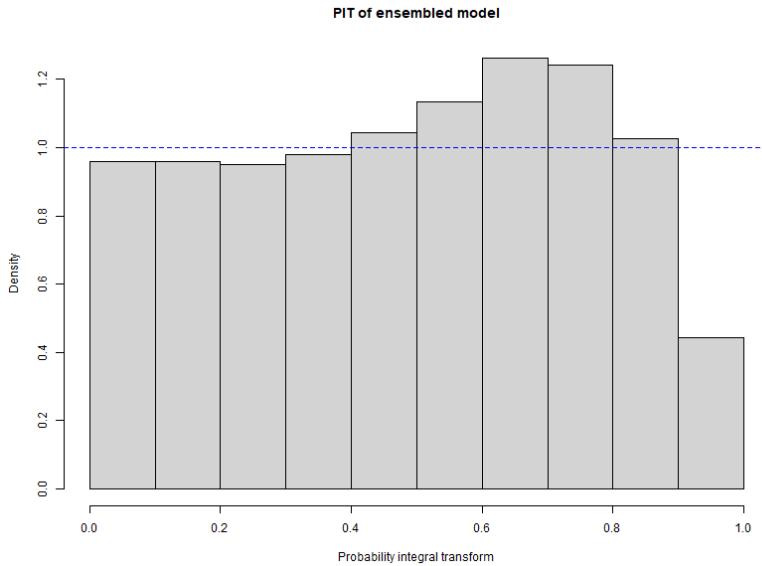


Figure 32: PIT-diagram of ensemble model

The PIT-diagram for the ensemble model in Figure 32 suggests that the model is generally well-calibrated. However, there are noticeable deviations. Specifically, observations of the most extreme avalanche counts are fewer than expected, while there are more observations of high but not maximal avalanche counts than would be anticipated if the model were perfectly calibrated.

Table 11: Performance of the ensemble model with random forest as the only explanatory variable.

RF ensemble model	
CRPS	0.11
RMSE	0.43

Table 11 shows the CRPS and the RMSE of the random forest ensemble model. These are very similar to the CRPS and RMSE of the ensemble model with multiple predictions as explanatory variables.

---

## 5.6 Comparison of the different model types

### 5.6.1 Using example dates

In this section, the forecasts of the models used in the ensemble model presented in Section 5.4.1 and the ensembled models on three different example dates are shown. Features of day 1, day 2 and day 3 are looked into in 4.6. Day 1 had 0 avalanches detected, Day 2 had 1 avalanche detected, while day 3 had 3 avalanches detected.

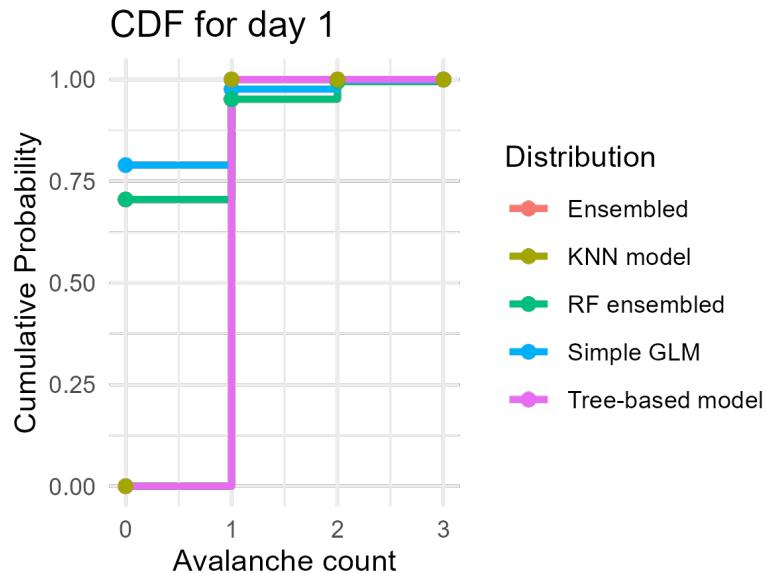


Figure 33: Cumulative distribution functions of model predictions on day 1.

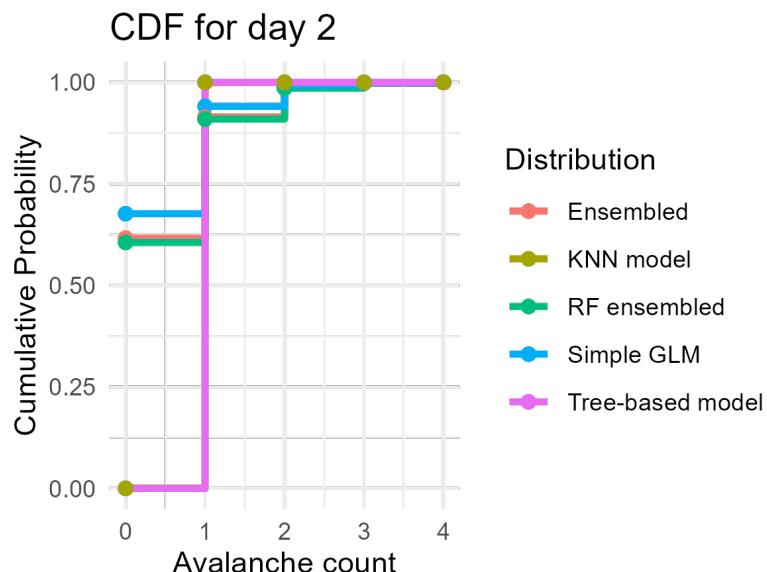


Figure 34: Cumulative distribution functions of model predictions on day 2.

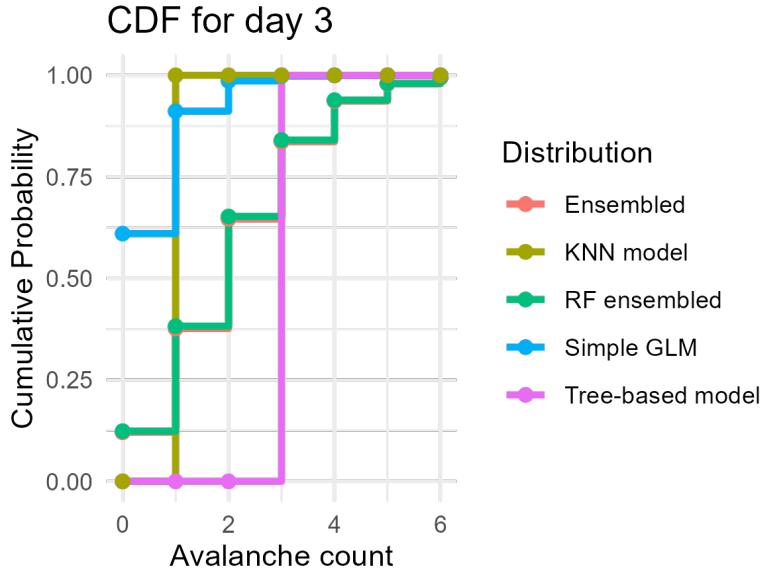


Figure 35: Cumulative distribution functions of model predictions on day 3.

The cumulative predictions of day 1 is presented in Figure 33. All of the models can with high confidence say that 1 avalanche or less is expected. The models fitted using the 2 and 5 nearest neighbours and  $\mathbf{X}_{\text{full}}$  and the model fitted using the 2 nearest neighbours and  $\mathbf{X}_{\text{reduced}}$  has the best predictions. They predict exactly 0 avalanches. The rest of the models predicts an avalanche count of less than 0.5, except the pruned tree fitted using  $\mathbf{X}_{\text{reduced}}$  predicting 1.154 avalanches and the generalized linear model fitted using interaction terms between the explanatory variables predicting 1.693 avalanches.

The cumulative predictions of day 2 is presented in Figure 34. These predictions are very similar to the predictions shown in Figure 33. The random forest model fitted using  $\mathbf{X}_{\text{full}}$  has the best prediction, it predicts 0.670 avalanches. If we round to the nearest integer, the random forest models and the pruned tree fitted using  $\mathbf{X}_{\text{full}}$ , the random forest ensemble model, and the models fitted using the 2 nearest neighbours would predict 1 avalanche. The rest of the models would predict an avalanche count different from 1.

The cumulative predictions of day 3 is presented in Figure 35. The models predicts an avalanche count of 3 or less with high confidence. The ensemble model has the best prediction with 2.11 avalanches. None of the models predicts an avalanche count of more than 2 if we round to the nearest integer.

Tables 13, 14, 17, 18, and 19 in the Appendix presents the predicted avalanche count from the different models.

---

### 5.6.2 Comparing models with baseline models

The model fits are compared to three baseline models. The baseline models are generalized linear models fitted with no explanatory variables, danger level as response, and observed avalanche count with lag 1.

Table 12: Performance of generalized linear model fitted using intercept-only and using danger level as explanatory variables.

	Danger level	Intercept-only	Lag 1 model
CRPS	<b>0.28</b>	0.31	0.30
RMSE	<b>0.87</b>	0.94	0.93

In Table 12 the performance of the generalized linear model fitted using only the intercept (which will predict the mean every time), the generalized linear model fitted using the danger level as explanatory variable and the generalized linear model with lag 1 avalanche count as explanatory variable are evaluated. It can be seen that the model fitted using the danger level as explanatory variable has the lowest CRPS and RMSE among the three models in Table 12. The performance of the two other are almost identical, meaning avalanche count from the day before gives little new information. This is also indicated by the weak correlation in Figure 11. Therefore, only the intercept-only model and the danger level model will be discussed.

The only models with both CRPS and RMSE lower than the generalized linear model fitted using the danger level as explanatory variable are the random forest fitted using  $\mathbf{X}_{\text{reduced}}$  in Table 4 and the ensemble models in Table 9 and 11. The random forests in Table 4, the simple generalized linear model in Table 7, and the ensemble models in Table 9 and 11 has lower or equal CRPS and RMSE than the intercept-only model. On the other side, the only model that performs worse than the intercept-only model with both a higher CRPS and RMSE is the generalized linear model fitted with interaction terms between the explanatory variables in Table 7.

---

## 6 Discussion and conclusion

### 6.1 Discussion

For a model to be useful and have potential for avalanche forecasting in Holmbuktura, it should at least perform better than a model which guesses the mean every day. The random forest models, the simple generalized linear model, and the ensemble models are performing better than the intercept-only model, and therefore they could all have potential as forecasts. On the other hand, there is already an existent forecast tool being used for avalanche forecasting. The regional avalanche warning issued as a danger level. A new model should therefore also perform better than a model using the danger level as input to be of interest for operational avalanche forecasting. The random forest model fitted using  $\mathbf{X}_{\text{reduced}}$  and the ensemble models also perform better than a model using danger level as input. Modelling avalanche count as a random forest seems to be the best modelling approach which are not a combination of forecasts. This is consistent with finding from Hennum (2016). However, the ensemble models have the lowest values of RMSE and CRPS. The two ensemble models have similar performance and PIT-diagrams, and hence it seems like there is little to gain by adding predictions from the generalized linear model and nearest neighbours model.

The most important variables for avalanche forecasting were found to be the day-ahead forecasts of new snow and precipitation. These variables are suggested as the most important variables by the simple regression trees, random forests, boosting models and the generalized linear models. The analysis shows that days with more precipitation and new snow are associated with an higher expected avalanche count, which aligns with the existing knowledge about avalanches summarized in Section 2.1. The results did not find any informative relation between the satellite count data from a close area and avalanche count.

The models have been fitted and evaluated on the same winter seasons. Ideally, the models would be tested on another winter season to check if the results are generalizable to other seasons besides the two seasons used for model fitting. A possible solution to this problem would have been to fit the models on one season, and tested them on another. However, this would lead to a weaker data basis and a model fit fitted using data from only one winter season.

Total avalanche count in an area where the avalanche could hit the road is used as response. However, the main goal of this research is to be able to forecast the avalanches hitting the road in Holmbuktura. An alternative would be to model the avalanche count hitting the road, but then the models would have to be fitted on only 15 avalanche observations. This leads to a need to translate the forecasted avalanche count in the area where avalanches could hit the road, to a forecast of avalanche count hitting the road. Figure B.3 in Appendix B shows an ordinary regression tree fitted to the data where only avalanches hitting the road are included. With such few events, it is not expected to get an informative model.

---

The aim of this research was to establish and validate a statistical method for a site-specific probabilistic forecast for avalanche releases, and This aim was pursued by collecting data from two winter seasons in Holmbuktura. The root mean squared error and mean continuous probability scores of the ensemble model indicates the ensemble models being the best models for avalanche forecasting. The histogram of the probability integral transforms shows the ensemble models also being well calibrated. Further, among the deterministic forecasts, the random forests performed well with low values of mean squared error and mean continuous probability scores. The random forest fitted using  $\mathbf{X}_{\text{reduced}}$  is also the most important forecast in the ensemble models.

The different evaluation criteria pointed at the ensemble models as the best models for forecasting avalanches. Both the ensemble models and the random forest fitted using  $\mathbf{X}_{\text{reduced}}$  performed better than all of the baseline models. Based on these results, fitting ensemble models and random forests could be a promising approach for operational avalanches and should be investigated further. In Section 2.3.4, probabilistic ensemble models as a combination of forecasts with different initial conditions are introduced. Further research could focus on fitting ensemble models as a combination of random forest predictions, by for example fitting multiple random forests, and by possibly varying the different initial conditions. Another alternative could be to estimate uncertainty by looking at the individual trees of the random forest models.

---

## Bibliography

- Amundsen, B. O. (2019, February 22). *Automatiske bommer skal stenge vegen når snøskredet går* [Veier24.no]. Retrieved 10th November 2023, from <https://www.veier24.no/artikler/automatiske-bommer-skal-stenge-vegen-nar-snoskredet-gar/458697>
- Atkinson, B. (n.d.). *Rpart*. Retrieved 12th May 2024, from <https://github.com/bethatkinston/rpart>
- Breiman, L., & Cutler, A. (n.d.). *RandomForest* [GitHub]. Retrieved 25th March 2024, from <https://github.com/cran/randomForest/blob/master/R/randomForest.default.R>
- Buser, O. (1983). Avalanche forecast with the method of nearest neighbours: An interactive approach. *Cold Regions Science and Technology*, 8(2), 155–163. [https://doi.org/10.1016/0165-232X\(83\)90006-X](https://doi.org/10.1016/0165-232X(83)90006-X)
- Czado, C., Gneiting, T., & Held, L. (2009). Predictive model assessment for count data. *Biometrics*, 65(4), 1254–1261. <https://doi.org/10.1111/j.1541-0420.2009.01191.x>
- Davis, R. E., Elder, K., Howlett, D., & Bouzaglou, E. (1999). Relating storm and weather factors to dry slab avalanche activity at alta, utah, and mammoth mountain, california, using classification and regression trees. *Cold Regions Science and Technology*, 30(1), 79–89. [https://doi.org/10.1016/S0165-232X\(99\)00032-4](https://doi.org/10.1016/S0165-232X(99)00032-4)
- Eckerstorfer, M., Malnes, E., & Müller, K. (2017). A complete snow avalanche activity record from a norwegian forecasting region using sentinel-1 satellite-radar data. *Cold Regions Science and Technology*, 144, 39–51. <https://doi.org/10.1016/j.coldregions.2017.08.004>
- Ellis, L. D., Badel, A. F., Chiang, M. L., Park, R. J.-Y., & Chiang, Y.-M. (2020). Toward electrochemical synthesis of cement—an electrolyzer-based process for decarbonating CaCO<sub>3</sub> while producing useful gas streams. *Proceedings of the National Academy of Sciences*, 117(23), 12584–12591. <https://doi.org/10.1073/pnas.1821673116>
- European avalanche warning services. (n.d.). *EAWS matrix*. Retrieved 24th November 2023, from <https://www.avalanches.org/standards/eaws-matrix/>
- Fahrmeir, L., Kneib, T., Lang, S., & Marx, B. (2013). *Regression: Models, methods and applications*. Springer. <https://doi.org/10.1007/978-3-642-34333-9>
- Gneiting, T., & Katzfuss, M. (2014). Probabilistic forecasting. *Annual Review of Statistics and Its Application*, 1(1), 125–151. <https://doi.org/10.1146/annurev-statistics-062713-085831>
- Gneiting, T., & Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477), 359–378. <https://doi.org/10.1198/016214506000001437>
- Haben, S., Voss, M., & Holderbaum, W. (2023). *Core concepts and methods in load forecasting: With applications in distribution networks*. Springer International Publishing. <https://doi.org/10.1007/978-3-031-27852-5>
- Hastie, T. (2009). *The elements of statistical learning: Data mining, inference, and prediction, second edition* (2nd ed.). Springer Nature.

- 
- Hennum, A. A. (2016). *Data-driven avalanche forecasting - using automatic weather stations to build a data-driven decision support system for avalanche forecasting* [Master thesis]. NTNU. Retrieved 23rd January 2024, from <https://ntnuopen.ntnu.no/ntnu-xmlui/handle/11250/2392436>
- Herwijnen, A. v., Mayer, S., Guillén, C. P., Techel, F., Hendrick, M., & Schweizer, J. (2023). DATA-DRIVEN MODELS USED IN OPERATIONAL AVALANCHE FORECASTING IN SWITZERLAND. *International Snow Science Workshop Proceedings 2023, Bend, Oregon*, 321–326. Retrieved 16th May 2024, from <https://arc.lib.montana.edu/snow-science/item.php?id=2895>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An introduction to statistical learning: With applications in r*. Springer US. <https://doi.org/10.1007/978-1-0716-1418-1>
- Krüger, F. (2023, December 11). *ScoringRules* [original-date: 2014-09-15T11:52:42Z]. Retrieved 13th March 2024, from <https://github.com/FK83/scoringRules>
- Kuhn, M. (2024, March 24). *Caret* [original-date: 2014-05-16T15:50:16Z]. Retrieved 25th March 2024, from <https://github.com/topepo/caret>
- Liboschik, T. (2023, January 11). *Tscount* [original-date: 2015-02-11T09:43:53Z]. Retrieved 14th March 2024, from <https://github.com/cran/tscount>
- Lied, K., & Kristensen, K. (2003). *Snøskred: Håndbok om snøskred*. Vett & viden I samarbeid med NGI, Norges geotekniske institutt. Retrieved 15th January 2024, from [https://urn.nb.no/URN:NBN:no-nb\\_digibok\\_2012051805067](https://urn.nb.no/URN:NBN:no-nb_digibok_2012051805067)
- Meier, L. (2018, July 5). *Radarmålinger av snøskred ved fv. 293 holmbuktura: Resultater fra testmålinger i 2017 og 2018* (Report). Statens vegvesen. Retrieved 16th January 2024, from <https://vegvesen.brage.unit.no/vegvesen-xmlui/handle/11250/2672819>
- Ridgeaway, G. (n.d.). *Gbm*. Retrieved 26th March 2024, from <https://github.com/gbm-developers/gbm>
- Schweizer, J., Bruce Jamieson, J., & Schneebeli, M. (2003). Snow avalanche formation. *Reviews of Geophysics*, 41(4). <https://doi.org/10.1029/2002RG000123>
- United Nations Department of Economic and Social Affairs. (n.d.-a). *The 17 goals*. Retrieved 15th September 2023, from <https://sdgs.un.org/goals>
- United Nations Department of Economic and Social Affairs. (n.d.-b). *Goal 9*. Retrieved 15th September 2023, from <https://sdgs.un.org/goals/goal9>
- Varsom. (n.d.-a). *Matrise for å sette faregrad*. Retrieved 24th November 2023, from <https://www.varsom.no/snokred/snokredskolen/del-informasjon/matrice-for-a-sette-faregrad/>
- Varsom. (n.d.-b). *Snøskredproblemer*. Retrieved 20th September 2023, from <https://varsom.no/snokred/snokredskolen/snokredproblemer/>
- Viallon-Galinier, L., Hagenmüller, P., & Eckert, N. (2023). Combining modelled snowpack stability with machine learning to predict avalanche activity [Publisher: Copernicus GmbH]. *The Cryosphere*, 17(6), 2245–2260. <https://doi.org/10.5194/tc-17-2245-2023>
- Widforss, A. (2023, September 26). *Regobslib* [original-date: 2021-08-23T08:31:04Z]. Norges vassdrags- og energidirektorat. Retrieved 21st February 2024, from <https://github.com/NVE/regobslib>

## Appendix

## A Data exploration

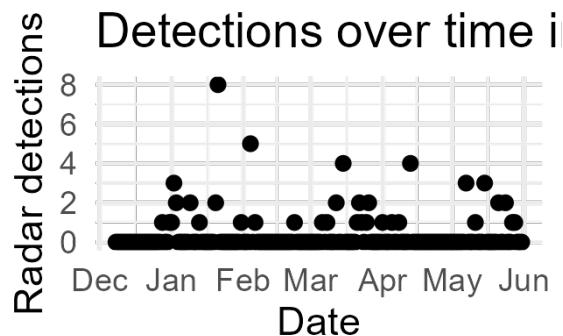


Figure A.1: Time series plot of number of avalanches detected by radar in Holmbuk-tura winter 2019/2020

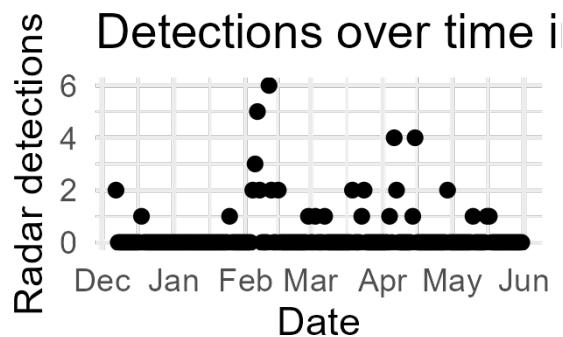


Figure A.2: Time series plot of number of avalanches detected by radar in Holmbuk-tura winter 2020/2021

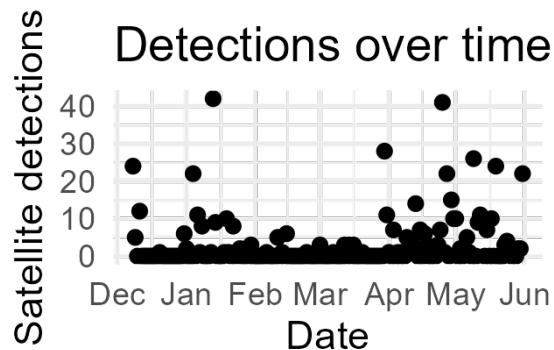


Figure A.3: Time series plot of number of avalanches detected by satellites in Lavangsdalen winter 2019/2020

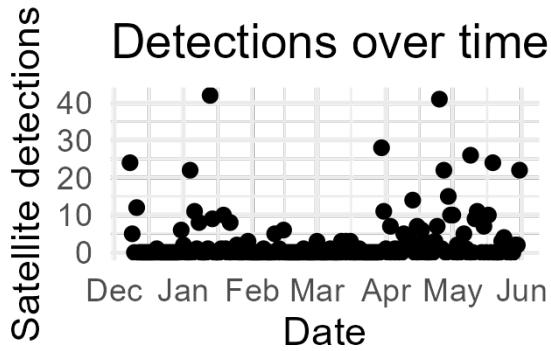


Figure A.4: Time series plot of number of avalanches detected by satellites in Lavangsdalen winter 2020/2021

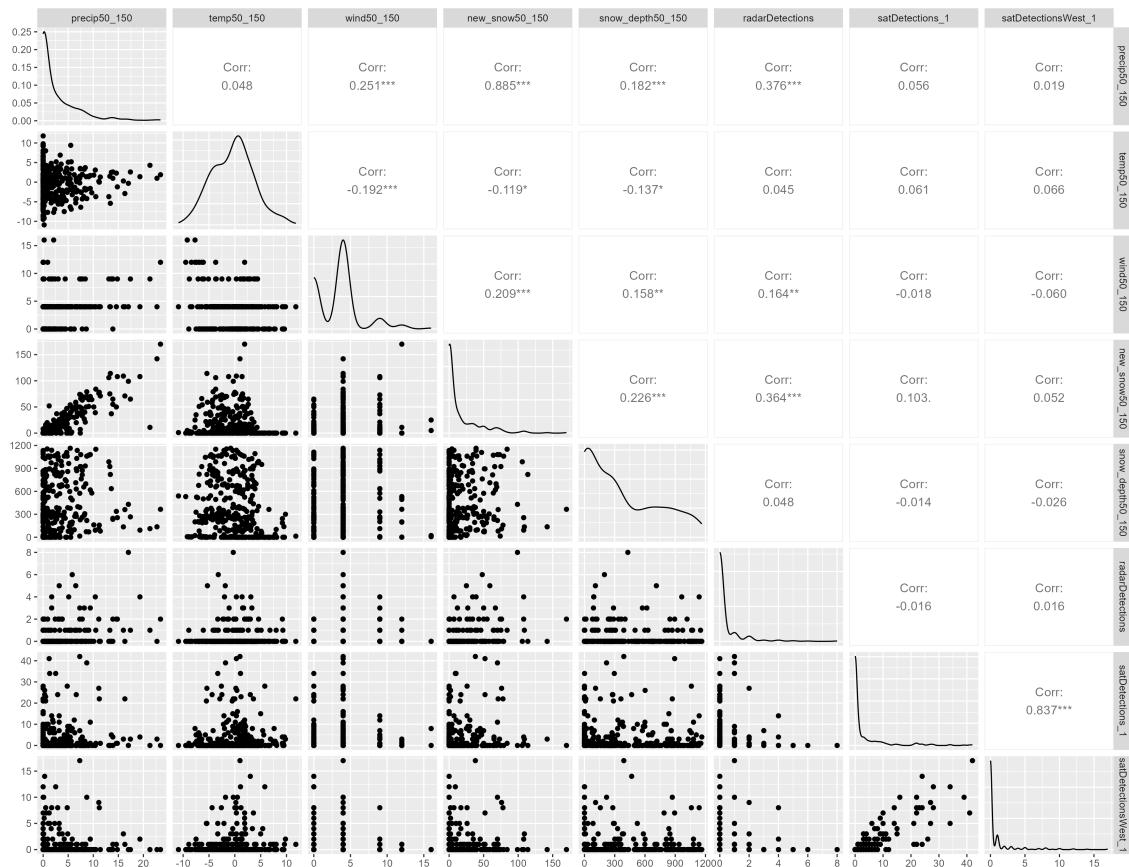


Figure A.5: Plot showing the correlation between a selection of variables. A star indicates the significance of the correlation. The 0.5 quantile at height 0-300 meter above sea level is used for the different weather forecasts.

## B Tree models

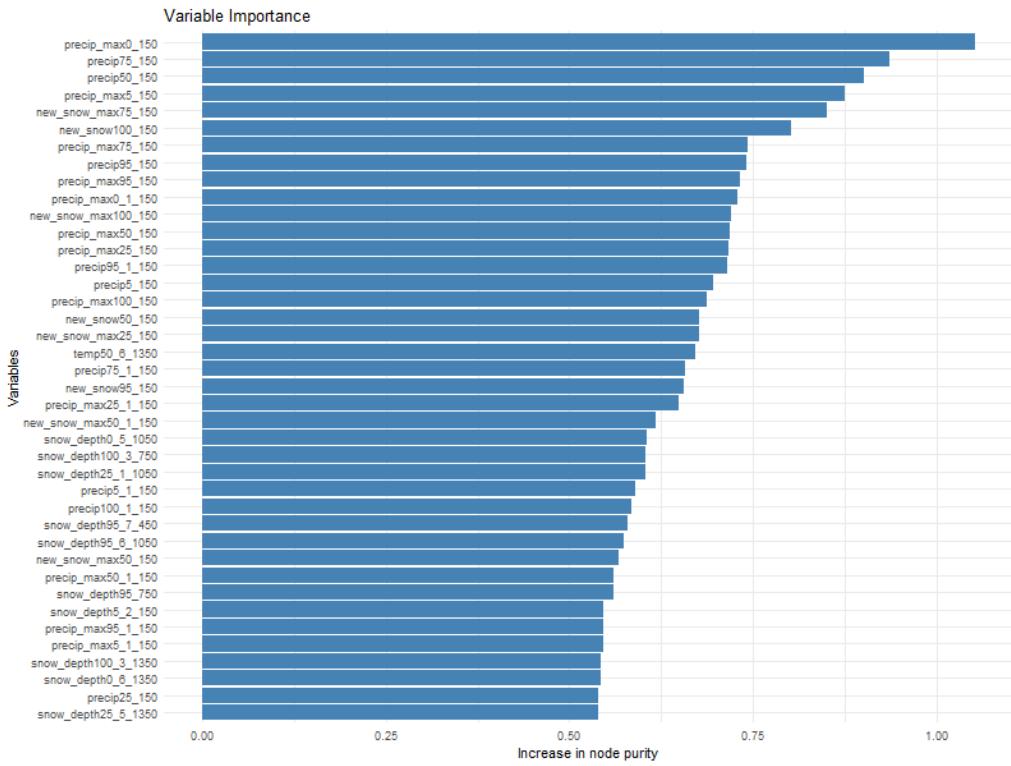


Figure B.1: Variable importance plot of the random forest model fitted using  $X_{\text{full}}$ .

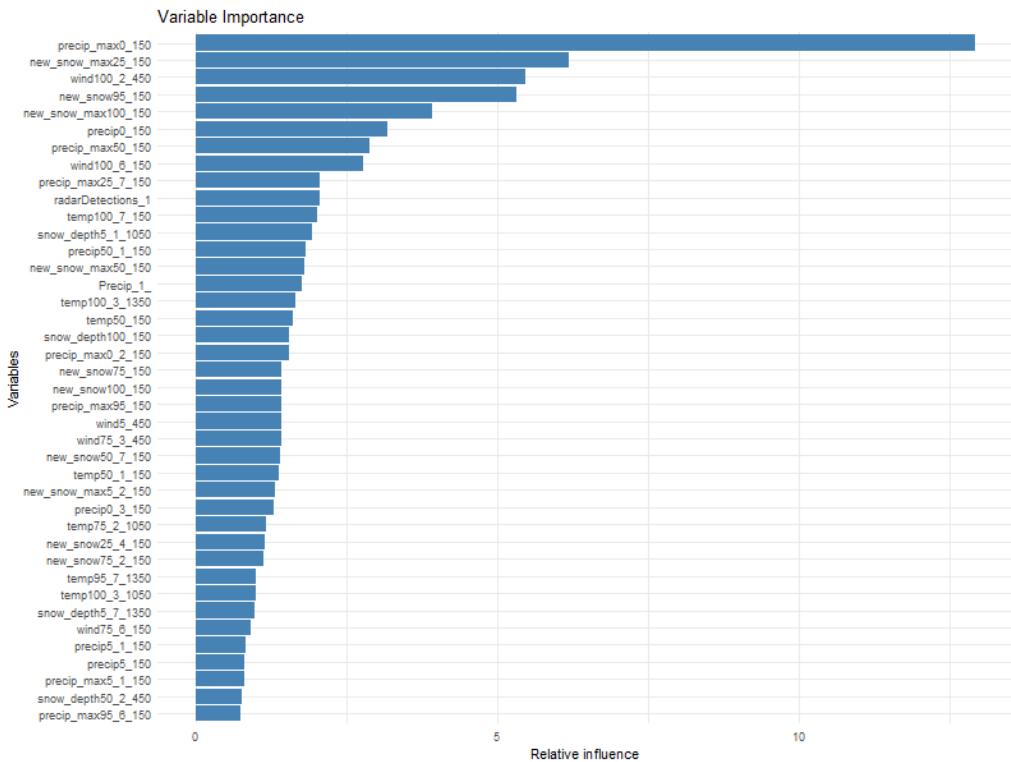


Figure B.2: Variable importance plot of the boosting model fitted using  $X_{\text{full}}$ .

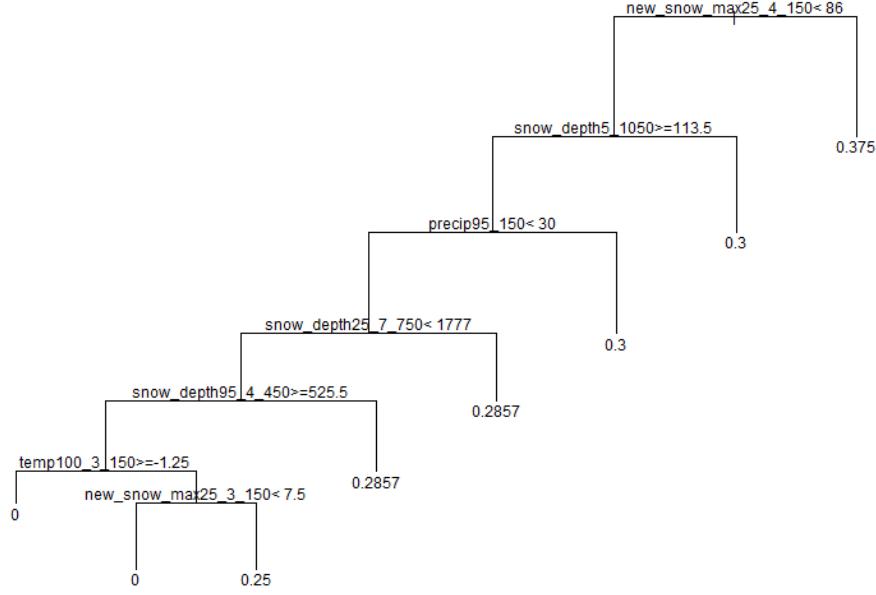


Figure B.3: Tree fitted using  $X_{\text{full}}$  and only avalanches hitting the road as response.

Table 13: Predictions from the various tree-based models on selected dates with different avalanche count detections.

	Pruned $X_{\text{full}}$	RF $X_{\text{full}}$	Boost $X_{\text{full}}$	Pruned $X_{\text{reduced}}$	RF $X_{\text{reduced}}$	Boost $X_{\text{reduced}}$
Day 1	0.08	0.23	0.12	1.15	0.35	0.20
Day 2	0.64	0.67	0.17	0.24	0.50	0.19
Day 3	1.43	1.82	1.66	1.15	2.09	1.08

## C Models fitted using the method of nearest neighbours

Table 14: Predictions from the various models fitted using the method of nearest neighbours on selected dates with different avalanche count detections.

	2NN $X_{\text{full}}$	5NN $X_{\text{full}}$	9NN $X_{\text{reduced}}$	2NN $X_{\text{reduced}}$	5NN $X_{\text{reduced}}$	20NN
Day 1	0.00	0.00	0.16	0.00	0.17	0.30
Day 2	0.50	0.33	0.37	0.50	0.17	0.15
Day 3	2.00	0.83	0.37	2.00	1.00	0.45

---

## D Generalized linear models

Table 15: P-values and estimates of the parameters of the generalized linear model with interaction terms.

	P-values	Coefficients
(Intercept)	0.34	0.31
precip_max0_150	$4.76 \times 10^{-3}$	0.22
new_snow100_150	0.3	$-2.90 \times 10^{-3}$
temp95_1_750	0.81	-0.01
wind100_750	0.19	-0.03
Precip_1	0.14	0.06
snow_depth25_750	0.11	$-4.15 \times 10^{-4}$
precip_max0_150:new_snow100_150	0.3	$1.13 \times 10^{-4}$
precip_max0_150:temp95_1_750	0.1	0.01
precip_max0_150:wind100_750	0.15	$-3.51 \times 10^{-3}$
precip_max0_150:Precip_1	0.82	$7.69 \times 10^{-4}$
precip_max0_150:snow_depth25_750	0.09	$-7.59 \times 10^{-5}$
new_snow100_150:temp95_1_750	0.04	$-4.05 \times 10^{-4}$
new_snow100_750:wind100_750	0.98	$2.86 \times 10^{-6}$
new_snow100_150:Precip_1	0.09	$-1.77 \times 10^{-4}$
new_snow100_150:snow_depth25_750	0.02	$3.26 \times 10^{-6}$
temp95_1_750:wind100_750	0.51	$-1.38 \times 10^{-3}$
temp95_1_750:Precip_1	0.05	-0.01
temp95_1_750:snow_depth25_750	0.06	$4.88 \times 10^{-5}$
wind100_750:Precip_1	0.9	$2.56 \times 10^{-4}$
wind100_750:snow_depth25_750	0.03	$4.00 \times 10^{-5}$
Precip_1:snow_depth25_750	0.01	$-5.86 \times 10^{-5}$

Table 16: P-values and estimates of the parameters of the generalized linear model with polynomial explanatory variables.

	P-values	Estimates
(Intercept)	$8.78 \times 10^{-31}$	-1.77
poly(precip_max0_150, 3)1	$3.81 \times 10^{-3}$	5.91
poly(precip_max0_150, 3)2	0.68	0.61
poly(precip_max0_150, 3)3	0.18	-1.83
poly(new_snow100_150, 3)1	$2.49 \times 10^{-6}$	10.58
poly(new_snow100_150, 3)2	$1.54 \times 10^{-4}$	-6.67
poly(new_snow100_150, 3)3	0.81	0.36
poly(temp95_1_750, 3)1	0.37	2.12
poly(temp95_1_750, 3)2	0.57	-1.58
poly(temp95_1_750, 3)3	$4.28 \times 10^{-4}$	9.43
poly(Precip_1, 3)1	0.74	0.78
poly(Precip_1, 3)2	0.28	-3.02
poly(Precip_1, 3)3	0.06	4.42
poly(snow_depth25_750, 3)1	0.18	-2.98
poly(snow_depth25_750, 3)2	0.02	4.88
poly(snow_depth25_750, 3)3	0.38	1.72
poly(wind100_750, 3)1	0.07	3.81
poly(wind100_750, 3)2	0.65	-0.97
poly(wind100_750, 3)3	0.91	-0.24

Table 17: Predictions from the various generalized linear models on selected dates with different avalanche count detection.

	Simple	With interaction terms	With polynomials
Day 1	0.24	1.69	0.35
Day 2	0.39	2.48	0.43
Day 3	0.49	2.28	1.55

Table 18: Predictions from of the ensemble model on the selected dates with different avalanche count detections.

Model	
Day 1	0.35
Day 2	0.48
Day 3	2.11

---

Table 19: Predictions from of the ensemble mode with random forest as the only explanatory variable on the selected dates with different avalanche count detections.

Model	
Day 1	0.35
Day 2	0.50
Day 3	2.09

---

## E Code

<https://github.com/BenSig00/Site-specific-probabilistic-forecast-for-avalanches.git>