

Avalanche warning system validation using satellite data

Kebene Amanuel Baissa

January 20, 2024

Abstract

This project focuses on assessing the accuracy of avalanche forecasts from Varsom by correlating them with actual avalanche events observed via satellite. The analysis is based on historical forecast data from Varsom and avalanche observations from the Copernicus Sentinel-1 satellite, explicitly focusing on the Lyngen region in Norway during three winter months in 2020. The avalanche data, comprising point data for avalanche locations, enables a detailed spatial analysis.

The study employs statistical methods like hit rate, precision, and Generalized linear models to ascertain the precision and effectiveness of Varsom's avalanche predictions. It systematically compares the issued warnings against actual avalanche occurrences in the designated area and time. Additionally, the project conducts spatial point process analysis to identify which spatial covariates most significantly influence avalanche occurrences.

The findings reveal a non-linear relationship between the forecasted danger levels and the number of actual avalanche occurrences, underscoring the complexity of predicting avalanche occurrences based solely on warning levels. Moreover, the analysis highlights the critical role of the slope in determining avalanche frequency, indicating that topographical factors are essential in understanding and forecasting avalanche events.

Contents

Abstract	iii
Contents	v
1 Introduction	1
2 Background	3
2.1 Snow avalanches	3
2.2 Avalanche warning service	3
2.3 Avalanche danger scale	4
2.4 Copernicus Sentinel-1	5
3 Teoretical bakground	7
3.1 Generalized Linear Models: Poisson Regression	7
3.1.1 Notation	7
3.1.2 Poisson regression	7
3.2 Spatial point process	8
3.2.1 Notation	8
3.2.2 Point process and point pattern	9
3.2.3 Point process model	11
4 Study area and data	13
4.1 Study region	13
4.2 Data	13
4.2.1 Avalanche data	15
4.2.2 Varsom warning data	17
4.2.3 The GIS data	17
4.2.4 Exploratory data analysis	18
5 Methods and Models	23
5.1 Comparing the frequency and degree of danger	23
5.2 A comparative analysis of forecasted terrain aspects and observed avalanche locations	24
5.3 Poisson linear regression models	25
5.4 Spatial point process model	26
5.5 Software	27
6 Results	29
6.1 Comparation of frequency and level of danger	29

6.2	A comparative analysis of forecasted terrain aspects and observed avalanche locations	30
6.3	Poisson regression models	31
6.4	Point process model	34
7	Discussion and Conclusion	37
	Bibliography	39

Chapter 1

Introduction

Snow avalanches, hereafter called avalanches, are rapid and destructive down-hill movements of a mass of snow, often mixed with rock, ice, soil, and other debris, on a mountainside or steep slope. With mountainous terrain and substantial winter snowfall during the winter months, Norway is a typical area where avalanches occur. The combination of significant snow accumulation, steep slopes, rapid temperature changes, and high winds determines the scope of the avalanche as described in Lied and Kristensen (2003).

Snow avalanches can be a dangerous natural threat in Norway, especially in hilly areas. Every winter, there are accidents of different sizes resulting from avalanches. Road networks are particularly affected, but power lines, residential areas, skiers, and others engaged in outdoor activities are at risk (Lied and Kristensen 2003). The biggest problem is the death of human beings and damage to infrastructures. To save those damages, we need a reasonable forecast of the avalanches. More accurate warnings can result in significant cost and time savings.

Regobs/Varsom is a system utilized by private individuals and official authorities, including the Norwegian Road Administration, which relies on it to determine road closures and openings to share and access data. It is provided by the Norwegian Water Resources and Energy Directorate (NVE) in collaboration with the Meteorological Institute (MET), the Norwegian Road Administration (Statens vegvesen, SVV), and Bane NOR. Varsom.no is a website used to provide warnings of natural hazards in Norway since January 2013. The website publishes warnings about avalanches, floods, landslides, and risks associated with driving on frozen water.

The European Avalanche Warning Services (EAWS) was tasked with identifying what affects local avalanche risks (EAWS 2024). They created a process and a reference chart, known as the EAWS Matrix, to help forecasters uniformly determine danger levels. The EAWS Matrix helps predict avalanche risks by examining snowpack stability, frequency of snowpack stability, and avalanche size.

Using the EAWS system, Varsom issues daily avalanche warnings. This warning's reliability and effectiveness are essential. Statistical methods can be employed to evaluate the accuracy and efficiency of these predictions. This evaluation can be systematically approached by comparing the issued warnings with the actual avalanche occurrences in the specified areas and time periods. Statistical metrics such as hit rate, precision, and Generalized linear models can be used to measure the accuracy and efficiency of the predictions.

The avalanche occurrence data is satellite-observed data recorded by the Copernicus Sentinel-1 satellite. In recent years, data from Synthetic Aperture Radar (SAR) obtained via the Sentinel-1 satellite has been increasingly utilized for detecting avalanches. The advantage of using SAR data lies in its capability to oversee extensive areas while remaining impervious to cloud cover and variations in lighting conditions (Kummervold et al. 2018).

The spatial representation of avalanche activity is necessary for effective local avalanche forecasting, as it provides vital insights into aspects such as the frequency, size, and conditions under which avalanches occur (Stoffel et al. 1998; Kummervold et al. 2018). Point process analysis plays a significant role in this regard, as it allows for an in-depth study of the locations of avalanches, revealing patterns in their intensity based on various factors, including geographical location and weather conditions.

This project aims to develop and present a method for analyzing and evaluating the accuracy of avalanche forecasts from Varsom by using historical warning data in conjunction with satellite observations of actual avalanche occurrences.

Chapter 2

Background

2.1 Snow avalanches

Avalanches travel in a specific direction from their site of origin, where they start and pick up speed, through a track where they can pick up additional snow, to the runout zone, when they slow down and dump the snow. The precise lines dividing these zones depend on elements, including the avalanche's magnitude and course. The topography and the amount of snow involved in the avalanche event affect the extent of the region covered by the avalanche's track and runout zone (Toft et al. 2023; Lied and Kristensen 2003).

The effect of terrain aspects on the number of triggered avalanches is a point of consideration. For example, sunny slopes can be dangerous because of the rapid warming and the high likelihood of the presence of buried sun crusts. However, aspects impact is shown to be minimal when the slope's effect is considered (David McClung and Peter A Schaerer 2006).

2.2 Avalanche warning service

In snowy, mountainous environments like Norway, avalanche forecasting is essential for hazard avoidance. Prior to 2013, Norway lacked a comprehensive public warning system. Instead, we had a restricted avalanche warning system focused on meteorological factors like strong winds or heavy snow (Engeset 2013).

The first avalanche warnings service for Norway was issued by the Norwegian Avalanche Center on January 14, 2013 (Engeset 2013). NVE is responsible for avalanche warning services, which is a collaboration between NVE, SVV, and the MET. This service follows European standards and is part of a wider EAWS.

The service provides regional predictions for high-risk regions on a five-level danger scale and notifications for the rest of the country when the danger level hits 4 or

5. They aim to deliver the most accurate avalanche threat information with their resources. Although the service's forecasts help assess avalanche hazards, they are not absolute assurances, as the threat might alter depending on the weather, snow conditions, topography, and duration of the snow cover.

The avalanche warning service creates daily forecasts, outlining each forecast area's danger rating, problems, and at-risk terrain. According to Varsom, the prediction for today is revised as needed before 10 AM, and forecasts for tomorrow and the day after that are posted before 4 PM. The warning's primary goal is to prevent avalanche-related loss of life and property.

The avalanche warning service relies heavily on the judgments of people who work as forecasters and observers. These judgments are based on weather observations and forecasts, snowpack observations, and observations of avalanches. To our understanding, no statistical algorithms are involved in this transition from observations to forecasts. The forecasters and observers are from different organizations. Some are from MET, SVV, and NVE. It is also possible for volunteers to register observations on a website called www.regobs.no.

2.3 Avalanche danger scale

The avalanche danger scale that is posted by Varsom is the highest expected scale for the most exposed part of the region for that day. This danger scale is the same for the rest of Europe. Europe adopted the European Avalanche Danger Scale in 1993, classifying avalanche threats into five categories: level 1 is low danger, level 2 is moderate danger, level 3 is considerable danger, level 4 is high danger, and level 5 indicates a highly high danger (Müller et al. 2022). The regions for which the danger scale is forecasted are at a minimum of 100km^2 . The danger scale is given from which avalanche problems are expected using the EAWS Matrix shown in Figure 2.1. The variables in the Matrix are shown in Table 2.2. The scales in the Matrix in Figure 2.2 are the newest version; the scales have different values in our dataset, which is from 2020. The variable scale is shown in Table 2.1. The European Avalanche Danger Scale (EADS) rates avalanche danger levels based on how stable the snow is, how often the snow stability changes, and how big an avalanche might be. Higher danger levels correspond to more unstable snow, more frequent instability, and larger avalanches. However, the terms used for describing snow stability and the likelihood of avalanches were not clearly defined, leading to different understandings among forecasters (D. McClung and P. A. Schaefer 2023).

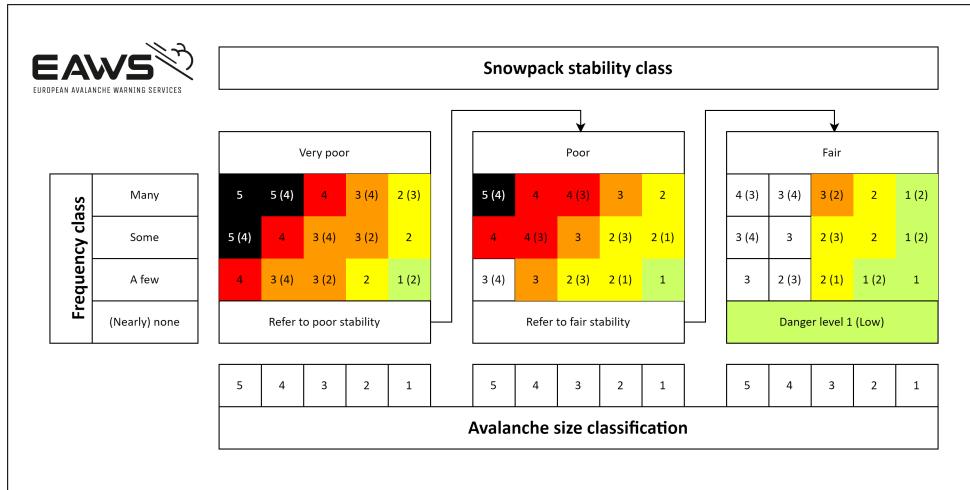


Figure 2.1: EAWS Matrix for objectively assessing the avalanche danger level (Version 06/2022) (Müller et al. 2022)

Table 2.1: Variables scale in the Varsom dataset

Variables	Scale from high to low risk in the dataset
Snowpack stability	45, 30, 20
Frequency class	Widespread, Specific, Isolated
Avalanche size	5,4,3,2,1

Table 2.2: Variables from the EAWS Matrix associated to the danger levels

Variables	Scale from high to low risk
Snowpack stability	Very poor, Poor, Fair
Frequency class	Many, Some, A few, (Nearly) none
Avalanche size	5,4,3,2,1

2.4 Copernicus Sentinel-1

The Copernicus Sentinel-1 mission consists of two satellite constellations (S1A and S1B). Each equipped with active radar sensors and microwave radiation instruments makes them able to offer continuous imagery of the Earth's surface, unaffected by weather and lighting conditions; however, the mission for Sentinel-1B ended in 2022 (ESA 2024).

Traveling in ascending and descending orbits, these satellites offer different revisit frequencies: about three days in equatorial regions and daily at high latitudes. Remote sensing for avalanche detection has limitations, notably in accurately assessing the probability of detection (POD) due to the disparity in dataset sizes

compared to real-world avalanches. This leads to uncertainty about the number of undetected avalanches. The average POD is 67.2%, alongside an average false alarm rate (FAR) of 45.9% (Eckerstorfer, Malnes et al. 2018; Eckerstorfer, Vickers et al. 2019).

Chapter 3

Teoretical bakground

3.1 Generalized Linear Models: Poisson Regression

3.1.1 Notation

We'll establish some essential notation in simple terms to lay the groundwork for the following theory.

In the provided notation, $\mathbf{y} = (y_1, \dots, y_{90})^T$ represents the number of avalanches observed each day in the Lyngen area over a specific period. Here, y_i is considered as a realization of the random variable Y_i . The subscript t denotes the particular day, with $t = 1$ corresponding to January 1, 2020, and $t = 90$ to March 30, 2020. The superscript T indicates that \mathbf{y} is a transposed vector, essentially converting it from a row to a column vector.

$\mathbf{X} = (\mathbf{X}_0, \mathbf{X}_1, \dots, \mathbf{X}_p)$ is used to denote a collection of covariates. This set of covariates is organized into a matrix of dimensions $90 \times (p+1)$, where $t = 90$ represents the total number of observations (one for each day in the specified period), and $p + 1$ represents the number of distinct covariates. Each row corresponds to a unique daily observation in this matrix format, and each column is associated with a specific covariate.

In the model, the coefficients linked with the covariates are represented by $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_p)^T$. β_0 serves as the intercept, aligning with the covariate vector $\mathbf{X}_0 = (1, \dots, 1)^T$. The linear predictor for each observation is defined as $\eta_i = \mathbf{x}_i \boldsymbol{\beta}$.

3.1.2 Poisson regression

The Poisson regression model theory below is based on the theory in the Book Fahrmeir et al. (2013).

The Poisson regression model is a widely used statistical approach designed for modeling count data. Specifically, it is adept at representing an event's frequency

of occurrence within a specified period. This model is particularly applicable when the variable of interest, Y , denotes the count of occurrences of a particular event.

The Poisson distribution for a random variable Y has the following probability mass function for a given value $Y = y$:

$$P(Y = y|\lambda) = \frac{e^{-\lambda}\lambda^y}{y!}, \quad (3.1)$$

for $y = 0, 1, 2, \dots$. The Poisson distribution is characterized by the parameter λ , which is the mean rate of occurrence for the event being measured and is a positive constant and the canonical link function, $\eta_i = \log(\lambda_i)$ or $\lambda_i = \exp(\eta_i)$. Resulting in

$$\lambda = \exp(X\beta). \quad (3.2)$$

Thus, the Poisson regression model for observation i can be written as:

$$P(Y_i = y_i|X_i, \beta) = \frac{e^{-\exp(X_i\beta)} \exp(X_i\beta)^{y_i}}{y_i!}. \quad (3.3)$$

For a sample size of n , the likelihood for a Poisson regression is given by:

$$L(\beta; y, X) = \prod_{i=1}^n \frac{e^{-\exp(X_i\beta)} \exp(X_i\beta)^{y_i}}{y_i!}. \quad (3.4)$$

The log-likelihood is

$$\ell(\beta) = \sum_{i=1}^n y_i X_i \beta - \exp(X_i \beta) - \sum_{i=1}^n \log(y_i!). \quad (3.5)$$

The maximum likelihood (or log-likelihood) function doesn't yield a straightforward analytical solution. Hence, an iterative method such as the iteratively re-weighted least squares is employed to approximate the values of the regression coefficients β .

3.2 Spatial point process

3.2.1 Notation

In this study, $W \subseteq \mathbb{R}^d$ represents the observation window, specifically focusing on the Lyngen region. The variable $n \in \mathbb{N}$ denotes a finite number of points, which in this context refers to the number of avalanches observed in the Lyngen region. The coordinates of these points are denoted by $\vec{x}_1, \dots, \vec{x}_n \in W$, representing the (X, Y) coordinates in UTM Zone 33. $M \in \mathbb{R}^{n \times p}$ are the marks associated with

these points, such as the size of each avalanche. Additionally, $\mathbf{X} \in \mathbb{R}^{n \times p}$ represents spatial p covariates like aspects, height, and terrain slope of where the avalanches are observed for the n points.

In a bounded set $B \subseteq \mathbb{R}^d$, the value $N(B)$ is a random variable representing the random number of points in B and $N(B) < \infty$. We denote the volume of $B \subseteq \mathbb{R}^d$ by $v(B)$. This corresponds to length (\mathbb{R}), area (\mathbb{R}^2) and volume (\mathbb{R}^3). Volume is a generic term that can be used for any d . For our data set with coordinate values of X and Y, we have $d=2$.

$E(N(B))$ denotes the mean number of points in B. The intensity of points in B is given by the mean:

$$\Lambda(B) = E(N(B)) \quad (3.6)$$

Λ is called the intensity measure.

Points come from a random process described by density function $\lambda(\vec{x})$ under some continuity conditions, which indicates the expected number of points at a specific location, \vec{x} , and gives us the intensity function:

$$\Lambda(B) = \int_B \lambda(\vec{x}) d\vec{x} \quad (3.7)$$

3.2.2 Point process and point pattern

The Point process theory below is based on the theory in the Book Illian et al. (2008).

The point data gives the location of objects within an area, avalanches observed in Lyngen in our case. In addition to the coordinates, we have marks associated with points, which are attributes of the points, such as avalanche type and size. Covariates are other additional data that can be linked to the points as explanatory variables, e.g., slope and aspect.

Point processes are a description of the point pattern. It helps us understand key aspects like the local intensity, how the points are distributed in space, how these patterns relate to other factors, and how different types of data points vary in their distribution across the area.

When points appear independently and randomly with a constant rate of occurrence $\lambda(\vec{x}) = \lambda$, this process is modeled by a homogeneous Poisson distribution, known as a Poisson point process. A homogeneous Poisson point process N on \mathbb{R}^d with intensity $\lambda > 0$ is characterized by $N(B) \sim \text{Poisson}(\lambda v(B))$ for all $B \subseteq \mathbb{R}^d$ so that $0 < v(B) < \infty$. And for disjoint $B_1, \dots, B_K \subseteq \mathbb{R}^d$, where $0 < v(B_k) < \infty$ for $k = 1, \dots, K$. $N(B_1), \dots, N(B_K)$ are independent random variables (independent scattering). The homogeneous Poisson point process describes complete spatial randomness.

Alternatively, if the points still emerge independently, but the rate of occurrence varies across different areas, this leads to an inhomogeneous point process, where $\lambda(\vec{x})$ changes across space. This is known as an inhomogeneous Poisson process. An inhomogeneous Poisson point process N on \mathbb{R}^d with intensity function $\lambda(\vec{x}) : \mathbb{R}^d \rightarrow [0, \infty)$ is characterized by $N(B) \sim \text{Poisson}(\nu_\lambda(B))$ for all $B \subseteq \mathbb{R}^d$ so that $0 < \nu_\lambda(B) < \infty$ whenever $\nu(B) < \infty$.

$$\nu_\lambda(B) = \int_B \lambda(\vec{x}) d\vec{x} \quad (3.8)$$

And for disjoint $B_1, \dots, B_K \subseteq \mathbb{R}^d$, where $0 < \nu(B_k) < \infty$ and $k = 1, \dots, K$, $N(B_1), \dots, N(B_K)$ are independent random variables. Here, we have spatially varying intensity and the independent point location.

A point pattern analysis examines the distribution of points across different dimensions, such as space, time, or others, focusing on the interdependence among these points. This analysis categorizes point distributions into three main types. The first is a random pattern; the points are distributed without any specific order or influence from surrounding points. The second is a repulsive pattern; points are evenly spaced, maintaining a consistent distance from each other. The third is a clustered pattern where points are grouped closely, often influenced by an external factor or covariate. This classification helps understand the dataset's underlying distribution and spatial relationships.

An estimated value obtained from analyzing a spatial point pattern dataset can be utilized for exploratory data analysis regarding the observed pattern (Ripley 1988). For a stationary point process N on \mathbb{R}^d with intensity $\lambda > 0$, the Ripley's second moment function $K(r)$ where $K : [0, \infty) \rightarrow [0, \infty)$ is given by:

$$K(r) = \frac{1}{\lambda(\vec{x}_0)} \mathbb{E}_0[|N(b(0, r) \setminus \{0\})|] \quad (3.9)$$

$b(0, r)$ is a ball with center in \vec{x}_0 with radius $r \geq 0$ and $\mathbb{E}_0[|N(b(0, r) \setminus \{0\})|]$ is expected number of points in the ball, given the point at \vec{x}_0 . In point pattern analysis, the L function linearizes the K function, simplifying its interpretation. L-function $L : [0, \infty) \rightarrow [0, \infty)$ is defined as:

$$L(r) = \left(\frac{K(r)}{b_d} \right)^{1/d}, \quad b_d = \nu(b(0, 1)). \quad (3.10)$$

Here $d = 2$, $L(r) = \sqrt{\frac{K(r)}{\pi}}$.

$L(r) = r$ at every r indicates complete spatial randomness at every distance. When the observed $L(r)$ value is larger than r (when the graph is above the line), it indicates clustering, and $L(r)$ less than r indicates repulsion.

3.2.3 Point process model

Point process model (PPM) is a Poisson point process model that examines the geometric arrangement of randomly distributed objects across one-, two-, or three-dimensional spaces. Point Process Models (PPMs) are similar to regression models like GLMs in Section 3.1.2, but their application differs. Standard regression models generally focus on a random variable Y_i , modeling its mean, μ_i , based on various covariates X_i while PPMs primarily focus on the spatial locations, \vec{x} , where observations are made, spotlighting the placement of these points. PPMs integrate these locations with the frequency of observed occurrences (n), modeling them by estimating the intensity, $\lambda(\vec{x})$, which is the expected number of observations per unit area. This intensity is then linked to regression techniques by modeling it with environmental and spatial covariates X observed throughout the study area W . (Renner et al. 2015) $\lambda(\vec{x})$ is modeled as the same in as shown in Equation 3.2, but adjusted for spatial dependence.

$$\lambda(\vec{x}) = X(\vec{x})\beta. \quad (3.11)$$

Chapter 4

Study area and data

This chapter will explore the study area and the relevant data from the time period to our case study. We will provide a detailed overview of the geographical and environmental characteristics of the region, as well as a comprehensive analysis of the data in the specific period.

4.1 Study region

In this project, we have chosen Lyngen as our case study area. Lyngen is an 813-square-kilometer municipality in Troms county in the far north of Norway, see Figure 4.1. It is well-known for the Lyngen Alps, a spectacular mountain range offering some of Norway's most beautiful landscapes. Lyngen has very long winters with much snowfall. The steep landscape of the Lyngen Alps and the surrounding hills significantly affect how often and intensely the avalanches occur. Many factors, like sudden changes in temperature, new snow, wind, and activities like hiking or skiing, can lead to these natural events. Varsom posts avalanche warnings for this region every day in the winter season. The area defined as Lyngen by Varsom extends beyond the boundaries of the Lyngen municipality. Figure 4.2 illustrates the Lyngen region as depicted by Varsom, and it is this specific area that we will focus on in our study.

4.2 Data

The data used for analysis in this project include the satellite-observed avalanche dataset, the Varsom warning dataset, and the altitude dataset. The avalanche dataset was downloaded as a CSV file from the NVE GitHub page. The warning dataset was obtained using an API in Python from API.NVE.NO. The altitude dataset is from hoydedata.no.

The altitude dataset is utilized to meticulously analyze the topography of the Lyn-



Figure 4.1: Lyngen in Norway's map

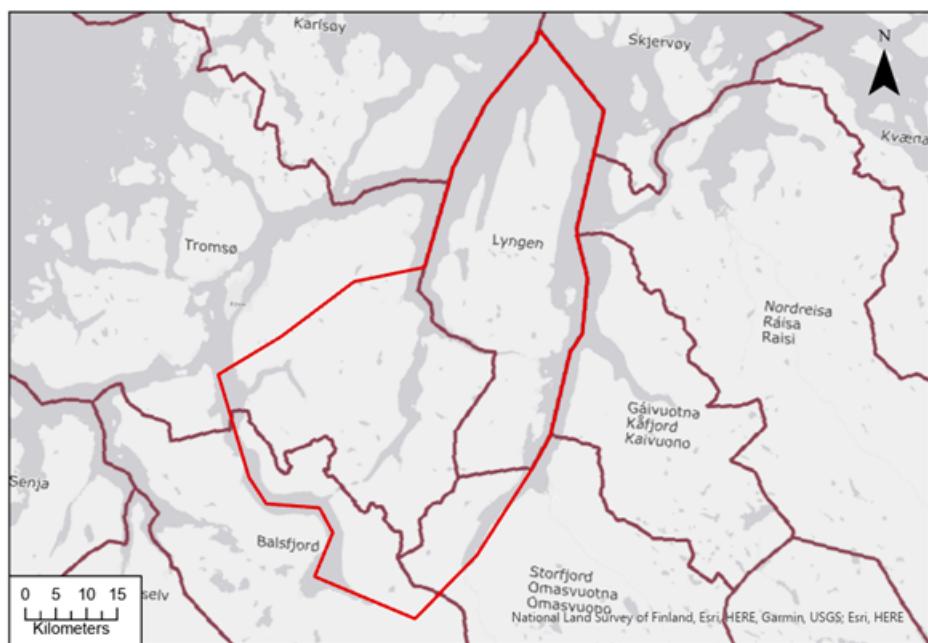


Figure 4.2: Region Lyngen as in Varsom

gen area using Geographic Information Systems (GIS). GIS provides information on the aspect and slope; see Figure 4.3, which shows the aspect for the specified area of Lyngen. Using a 100x100 raster grid format allowed us to evaluate the

slope, elevation, and aspect at specific points throughout the area while effectively managing data size and complexity. All the data and analyses align with the UTM Zone 33 coordinate system, which provides a standardized spatial referencing and mapping framework. Three datasets containing information on slope, elevation, and aspect corresponding to X and Y coordinates are downloaded from the GIS as Excel files.

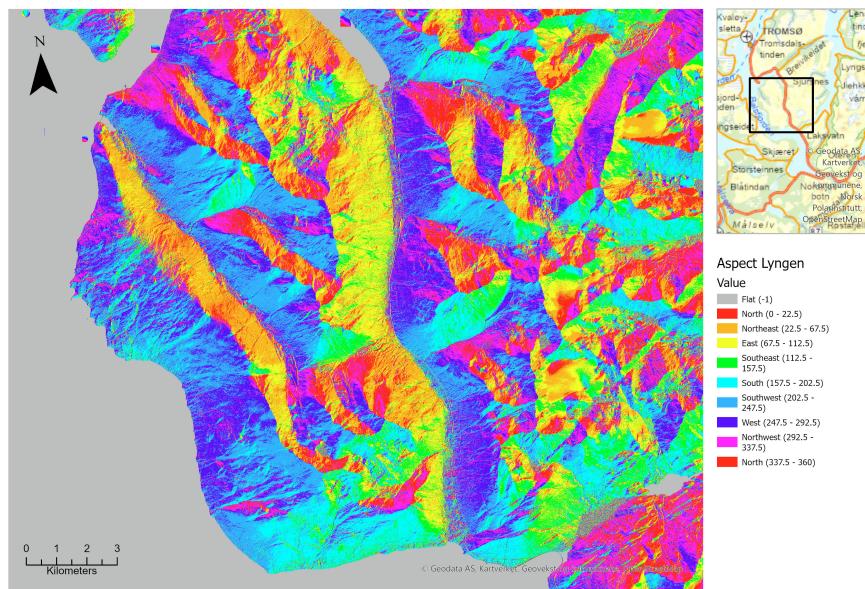


Figure 4.3: Aspect from GIS

4.2.1 Avalanche data

From January 1 to March 30, 2020, the Sentinel-1 satellite recorded 2247 avalanches in the Lyngen region, occurring over 61 of the 90 days, see Figure 4.4. The size of the circles represents the relative area of the avalanches observed.

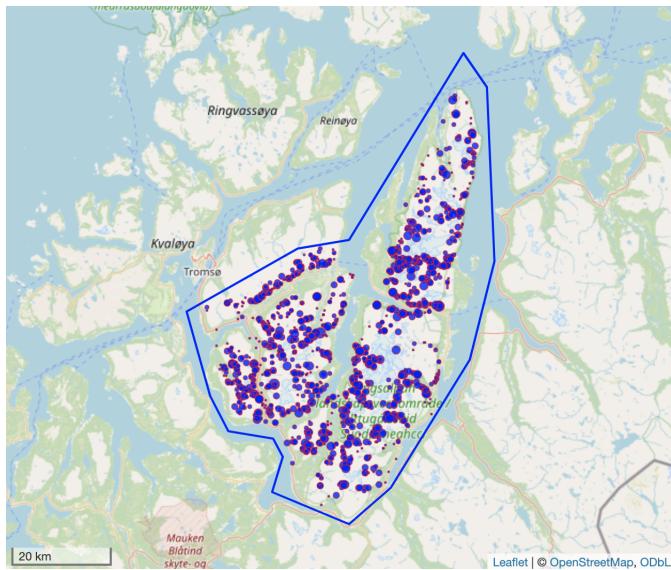


Figure 4.4: Avalanches recorded in Lyngen in winter 2020.

The avalanche data includes information only for the days when avalanches were observed; we do not have any data for 29 days out of the 90. We assume that no avalanches occurred these days, although there might be detection errors from the satellite. The number of avalanches observed each day varies greatly, with a minimum of 1 and a maximum of 390 on March 30, as shown in Figure 4.5. The avalanche runout aspects are measured in degrees, ranging from -180 to 180 in the dataset. These values are converted into directional characters such as N, NE, E, SE, S, SW, W, and NW, using the scale depicted in Figure 4.3 to make it comparable with the aspects in the warning dataset. The overview of the important variables of the avalanche dataset is shown in Table 4.1.

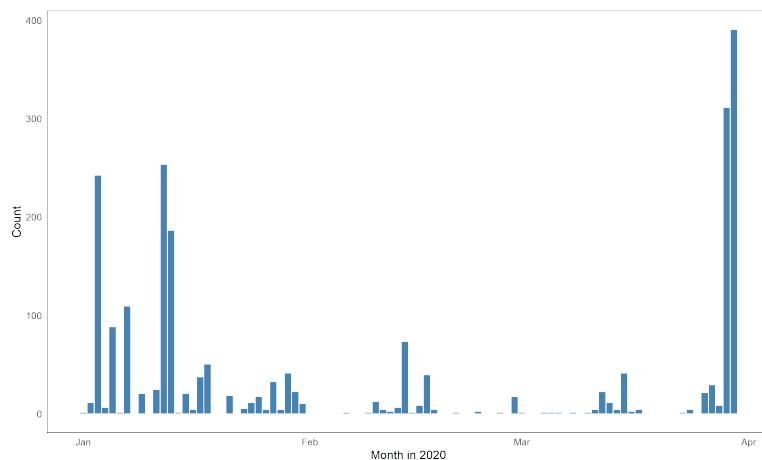


Figure 4.5: Distribution of the Sentinel-1 satellite detected avalanche data.

Table 4.1: A summary of the avalanche dataset

Name	Description	Value
SkredTidspunkt	Time of when the avalanche occurred	Date(yyyy-mm-dd)
Tidsnøyaktighet	Time accuracy	some days or hours
Kvalitet	Quality of the detection	Quality type A, B or C
X	Coordinate X	Eg. 700795.5
Y	Coordinate Y	Eg. 7736651
EksposisjonUtlopsomr	Aspect	(N,NE, E, SE, S, SW, W, NW)
SnittHelningUtlopsomr_gr	Mean angel of the runout	Slope in degree
MaksHelningUtlopsomr_gr	Maks angel of the runout	Slope in degree
MinHelningUtlopsomr_gr	Min angel of the runout	Slope in degree
Hoyde	Altitude	Height above sea level in meters
SHAPE_Length	Length of the avalanche	Length in meters
SHAPE_Area	Area of the avalanche	Area in square meters
Count	Amount of the avalanche got detected	1 for observed

4.2.2 Varsom warning data

The Varsom warning data contains the daily forecast from 01-01-2020 to 30-03-2020. This dataset contains the variables described in Table 4.2. Our analysis focuses on SP 1 (Skred problem 1), denoting the primary avalanche problem identified as the foremost cause of concern. Additionally, the data includes SP 2, representing the secondary avalanche problem. However, our attention will be primarily directed toward SP 1, as it is the principal factor in our study. An avalanche problem refers to specific conditions in the snowpack that can lead to avalanches. A summary of the dataset is given in Table 4.3.

4.2.3 The GIS data

The GIS datasets contain the slope, aspect, and elevation values at 450,269 points in the Lyngen area. To make the data more consistent with the datasets mentioned above, the aspects have been converted to directional characters such as N, NE, E, SE, S, SW, W, NW, and F, based on the scale shown in the map Figure 4.3. The slope values have also been converted to character representations, as shown in Figure 4.3, to make them suitable for analysis. Figure 4.6 visually represents the datasets.

Table 4.2: A summary of the Varsom warning dataset

Name	Description	Value
Dato	The date of the avalanche warning	Date(yyyy-mm-dd)
Varslingsregioner	Region	LYNGEN
Region_ID	Region id	3010
Faregrad	Danger scale	2, 3, 4
SP1	Avalanche problem	NEW_SLAB, WET_SLAB, WIND_SLAB
Utbredelse_SP1	Frequency class	SPECIFIC, WIDE-SPREAD, ISOLATED
Skredstorrelse_SP1	Avalanche size classification	2 , 3
Utlosbarhet_SP1	Snowpack stability class	20, 30, 45
Himmelretninger_SP1	Sky directions(aspect)	(N,NE, E, SE, S, SW, W, NW)

Table 4.3: Data summary from the historical warning data set

Danger	Count	Avg. Stab.	Avg. Size	Freq. Count		
				Iso.	Spec.	Wide.
2	51	27.65	2.08	4	45	2
3	38	31.18	2.66	2	25	11
4	1	30.00	3.00	0	0	1

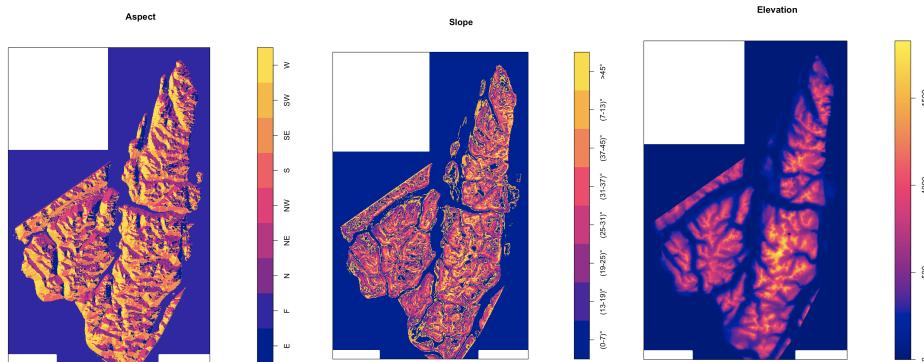


Figure 4.6: Terrain map

4.2.4 Exploratory data analysis

Once we've processed our dataset, we thoroughly examined the data using some straightforward statistical plots. These include pairs plots, mosaic graphs, and point pattern plots, each offering unique insights.

Pairs plots consist of a grid of scatterplots, each showing the relationship between two variables in our dataset. This setup is beneficial for spotting how various ele-

ments are correlated. Mosaic graphs are another tool we use. These are great for displaying relationships in the dataset, especially when comparing different groups or categories. They make it easier to see the big picture. Point pattern plots on maps are invaluable for datasets with a geographical element. Placing data points on a map allows us to see patterns and distributions tied to specific locations, offering insights into geographical sites. This combination of plots gives us a comprehensive view of our dataset, helping us understand the data from different angles.

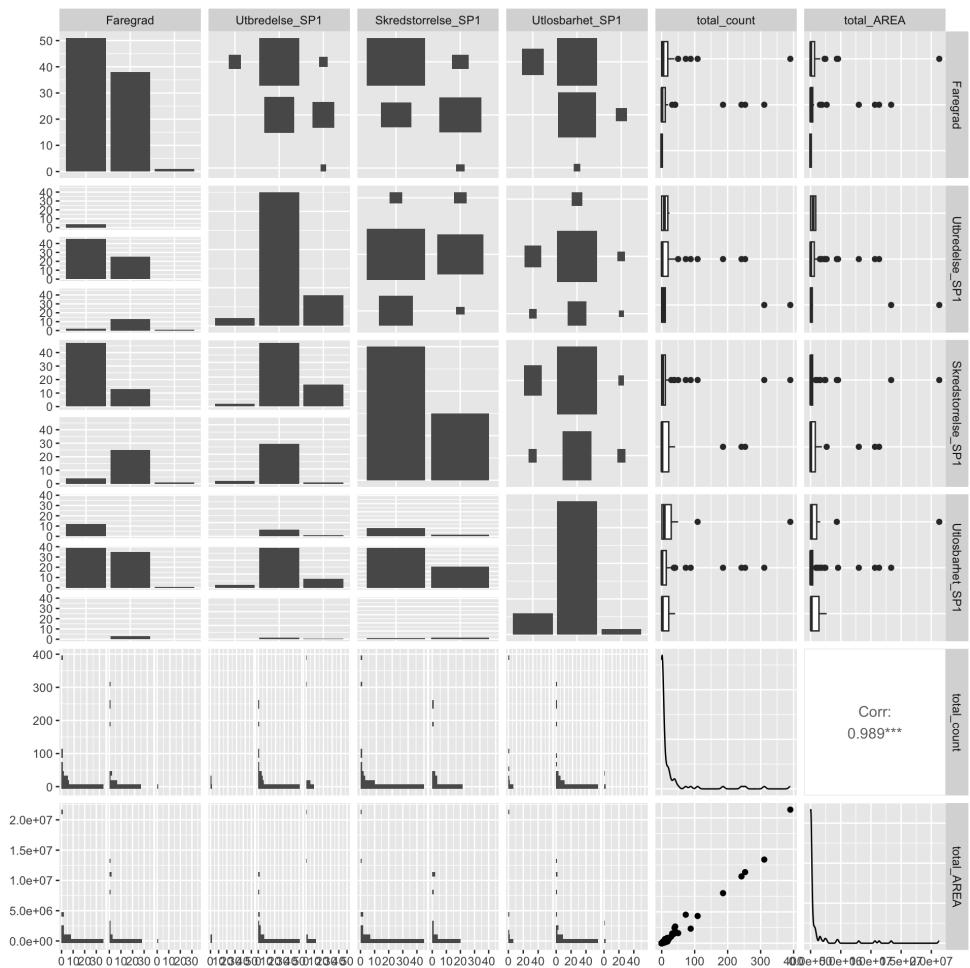


Figure 4.7: Pairs plot of the predicted value and observed values

The pairs plot Figure 4.7 shows a correlation between the variables. Snowpack stability (Utlosbarhet), Frequency class (Utbredelse), Avalanche class (Skredstørrelse), and the Level of danger (Faregrad). This proves that the Level of danger is the function of these variables. We can observe that the Level of danger is associated with the variables from the Mosaic plots in Figure 4.8. For example, most of the days when the Level of danger was 2 had predicted Avalanche Size 2, and

the days when the predicted value was 3 had Avalanche size 3. And the day with Level of danger 4 had 3 as Avalanche size.

From the pairs plot Figure 4.7 we can't see that the Level of danger affects Total count or Total area that much. We will study the direct correlation of the Total count and Level of danger in the model Section 5.1. The pair plot also shows a high correlation between the Total area of the avalanches observed and the Total count of the avalanches. Because of this, in this project, we focus only on the Total count.

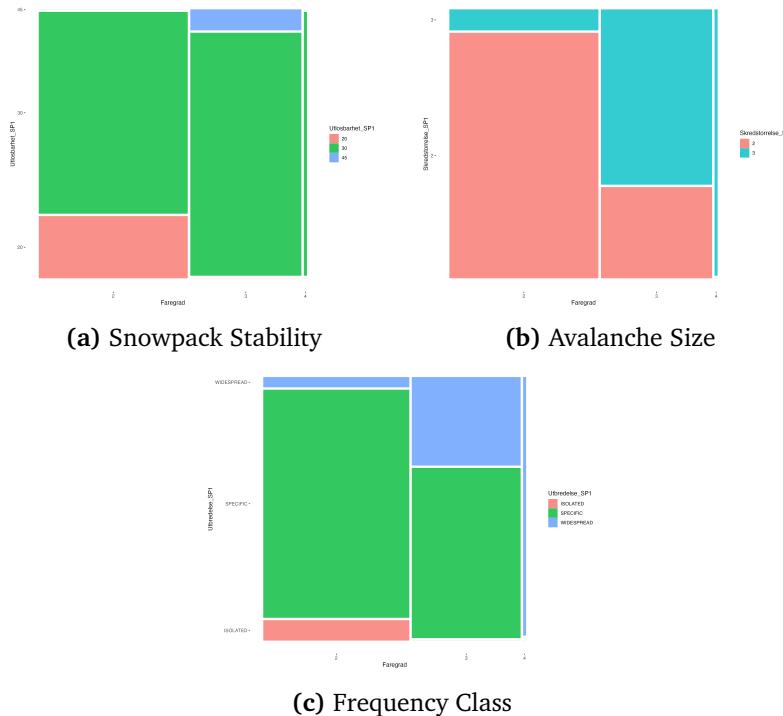


Figure 4.8: Mosaic Plot Showing the Association Between Degree of Danger Level and the Three Different Variables

Exploratory data analysis involves analyzing point patterns to understand and interpret the underlying spatial distribution. Using the L function described in Section 3.2.2 we can study the pattern and location of the points. Since there are a total of 2,247 points (avalanche observed) in the period from January 1, 2020, to March 30, 2020, which is a large number for studying the pattern of the points, we will focus on the first three days of January, which comprise 254 points, see Figure 4.9.

In R, the L function can be found using the function `Lest()` from the library `spatstat`.

If a stationary Poisson RF is a suitable model, estimated L-function lines should overlap the blue theoretical (Pois) line. The observed L(r) function in Figure 4.10

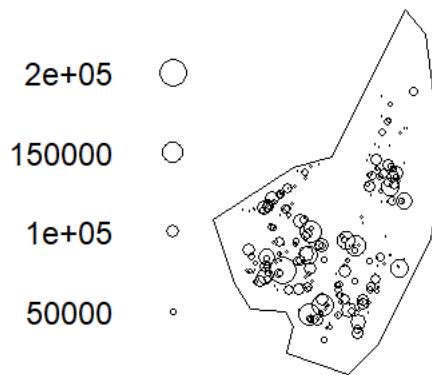


Figure 4.9: Avalanche observed from 01-01-2020 to 03-01-2020. Marked by their area size in square meters.

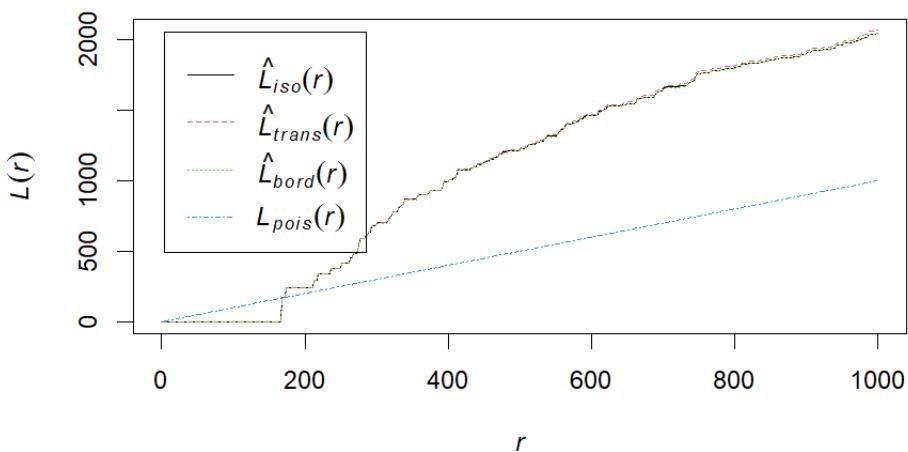


Figure 4.10: Empirical L-interaction function plotted against the theoretical with different edge corrections ("border", "isotropic", "Ripley", "translate").

is under the theoretical line, which indicates that the point distribution is more repulsive than a random distribution at a distance of under 175 meters to each other. The pattern is can also be seen in the Figure 4.9.

Chapter 5

Methods and Models

5.1 Comparing the frequency and degree of danger

When evaluating avalanche warnings during the winter period of 2020 compared to satellite-detected avalanches, we start to compare the count of avalanches that occurred versus the predicted warnings for each respective day. The degree of danger mainly falls into categories 2 and 3, with one instance of category 4. However, no registered avalanche occurred on the day when a category 4 warning was issued. We compare the degrees of 2 and 3 by analyzing each group's median, spread, and outliers. Side-by-side boxplots are used for each category to facilitate this comparison.

Boxplots summarize our data visually. They show us how the data is spread out and point out any outliers. The benefit of boxplots lies in their ability to compare different sets of data side by side in a single graph.

Hypothesis test

In hypothesis testing, data are used to infer which of two competing hypotheses, the null or alternative hypothesis, is correct. The null hypothesis (H_0) is that there is no difference in mean outcomes across the two different levels of danger. The alternative hypothesis (H_A) is that there is a difference in the means of the two groups when the degree of danger is 2 and 3. μ_2 is the mean of the number of avalanches that occurred when the danger level was 2, and μ_3 is the mean of the number of avalanches that occurred when the danger level was 3.

$$H_0 : \mu_2 = \mu_3 \quad H_A : \mu_2 \neq \mu_3$$

Anova model can be used to answer whether the groups' means differ from each other. ANOVA is abrivation for Analysis of variance. Anova test can be done by running ANOVA in R. 'aov()' function can be used.

5.2 A comparative analysis of forecasted terrain aspects and observed avalanche locations

To analyze the precision of the aspects predicted by Varsom, we calculate the ratio of the total count of avalanches that occurred on predicted aspects to the total observed avalanches for each day an avalanche occurs. By averaging these ratios across all days, we obtain a mean accuracy that reflects the overall predictive performance of the model. We can also compute statistical summaries such as the median, minimum, maximum, and standard deviation of these daily accuracy ratios to understand the model's consistency and reliability better.

Let O_t represent the set of observed avalanches on day t and P_t represent the set of predicted aspects for day t . The daily accuracy ratio, A_t , for each day t is given by:

$$A_t = \frac{|O_t \cap P_t|}{|O_t|} \quad (5.1)$$

Where:

$|O_t \cap P_t|$ is the count of observed avalanches that were correctly predicted on day t .

$|O_t|$ is the total count of observed avalanches on day t .

We define T as the set of all days when avalanches were observed. To get the summary, we can find the mean. The mean accuracy ratio across all days is:

$$\bar{A} = \frac{1}{|T|} \sum_{t \in T} A_t \quad (5.2)$$

Statistical summaries of the daily accuracy ratios can be expressed as:

- Mean accuracy ratio: \bar{A}
- Median accuracy ratio: $\text{median}(\{A_t \mid t \in T\})$
- Minimum accuracy ratio: $\min(\{A_t \mid t \in T\})$
- Maximum accuracy ratio: $\max(\{A_t \mid t \in T\})$
- Standard deviation of accuracy ratios: $\sqrt{\frac{1}{|T|-1} \sum_{t \in T} (A_t - \bar{A})^2}$

These statistical summaries provide a comprehensive evaluation of the predicted aspect in the warning data by indicating its average accuracy and the variability of accuracy from day to day.

We can also calculate the ratio of how many times each aspect is predicted over the period of 90 days. Let A represent the set of all possible aspects and $P(a, t)$ represent a prediction function for an aspect a on day t , which is 1 if the aspect

is predicted on that day, and 0 otherwise. The ratio $R(a)$ of predictions for each aspect a over 90 days is given by:

$$R(a) = \frac{1}{90} \sum_{d=1}^{90} P(a, d) \quad (5.3)$$

where $R(a)$ is the ratio of predictions for aspect a , and t ranges from 1 to 90, representing each day in the period.

The $R(a)$ will be compared with the estimate of the aspect covariates in the Point process model in Section 6.4.

5.3 Poisson linear regression models

The avalanche danger level is a function of snowpack stability, frequency of the avalanche, and avalanche size (Müller et al. 2022). Assuming the total count of avalanches each day to be Poisson distributed as in Equation 3.1, such that:

$$y_t \sim \text{Poisson}(\lambda_t) \quad (5.4)$$

The daily mean λ_t is linked to the explanatory variables or the covariates through a logarithmic link function. Since the avalanche danger level is a function of snowpack stability, frequency of the avalanche, and avalanche size, we use the variables as explanatory variables to study how they affect the total count of avalanches for each day. We get:

$$\text{Model 1 : } \eta_t = \beta_0 + \beta_1 p_t + \beta_2 f_t + \beta_3 s_t \quad (5.5)$$

The covariates are described in Table 5.1.

Table 5.1: Covariates used in Poisson model

Covariate	Description
p_t	Snowpack stability (Utlosbarhet)
f_t	Frequency class (Utbredelse)
s_t	Avalanche size (Skredstorrelse)

Model 2, includes an interaction term between Snowpack stability and Avalanche size, in addition to the main effects. Including the interaction term allows the model to assess whether the combined effect of these two variables on Total count differs from the sum of their individual effects.

$$\text{Model 2 : } \eta_t = \beta_0 + \beta_1 p_t + \beta_2 f_t + \beta_3 s_t + \beta_4 (p_t \times s_t) \quad (5.6)$$

Model 3 adds the interaction term between Frequency class and Avalanche size to Model 2.

$$\text{Model 3 : } \eta_t = \beta_0 + \beta_1 p_t + \beta_2 f_t + \beta_3 s_t + \beta_4 (f_t \times s_t) + \beta_5 (p_t \times s_t) \quad (5.7)$$

We use the AIC values to choose from the models. The Akaike Information Criterion (AIC) is a prevalent metric in likelihood-based inference, primarily used for model selection. It serves as a tool for comparing and selecting among various competing statistical models that differ in their predictors and parameters. The AIC seeks a balance between closely fitting the data and maintaining a reasonable level of model complexity, essentially avoiding excessive parameterization (Fahrmeir et al. 2013). This value can be found in the summary output of the model in R. We use the `glm()` function for the models in R.

5.4 Spatial point process model

The recorded avalanche occurrences in the avalanche data can be analyzed using a spatial point process model (PPM). These models are instrumental in determining the intensity of avalanches at various locations. Essentially, this intensity quantifies the expected number of occurrences at a specific location in space, providing a detailed understanding of avalanches' spatial distribution and frequency. By applying the inhomogeneous Poisson model with the intensity of the avalanches as a log-linear function of the slope, aspect, and elevation, the relationship between the environmental factors and the point process's intensity is studied. The model is:

$$\ln(\lambda(\vec{x})) = \beta_0 + A(\vec{x})\beta_1 + S(\vec{x})\beta_2 + E(\vec{x})\beta_3 \quad (5.8)$$

$A(\vec{x})$ is the terrain aspect at position \vec{x} in degrees, $S(\vec{x})$ is the terrain slope at position \vec{x} and $E(\vec{x})$ is the terrain elevation in meters. In the model, the elevation is scaled by 500 meters to make it more comparable with the other covariates. For this model, we use the covariates as categorical variables as described in Subsection 4.2; this gives:

$$\begin{aligned} \ln(\lambda) = & \beta_0 + A_F\beta_1 + A_N\beta_2 + A_{NE}\beta_3 + A_{NW}\beta_4 + A_S\beta_5 + A_{SE}\beta_6 + A_{SW}\beta_7 + A_W\beta_8 + \\ & S_{13-19}\beta_9 + S_{19-25}\beta_{10} + S_{25-31}\beta_{11} + S_{31-37}\beta_{12} + S_{37-45}\beta_{13} + S_{7-13}\beta_{14} + \\ & S_{>45}\beta_{15} + E\beta_{16} \end{aligned} \quad (5.9)$$

Table 5.2: Description of Predictors and Coefficients in the PPM

Symbol	Description
$\ln(\lambda)$	Natural logarithm of the intensity λ
β_0	Intercept term
$A_F, A_N, A_{NE}, A_{NW}, A_S, A_{SE}, A_{SW}, A_W$	Dummy variables for different aspects (directions), with coefficients $\beta_1, \beta_2, \dots, \beta_8$
$S_{13-19}, S_{19-25}, S_{25-31}, S_{31-37}, S_{37-45}, S_{7-13}, S_{>45}$	Variables for different slope categories, with coefficients $\beta_9, \beta_{10}, \dots, \beta_{15}$
$E(\vec{x})$	Elevation, multiplied by its coefficient β_{16}

We use the `ppm()` function for the model in R.

5.5 Software

In this project, Python (Rossum 1995) is used to access the API data of the warning dataset. ArcGIS was employed to extract terrain slope, aspect, and elevation information from a 100x100 raster of the study area.

The statistical models are implemented using R (R Core Team 2021). The packages used in R for handling geospatial data include the `sf` package (Pebesma 2018) for spatial data operations, the `leaflet` package (Cheng et al. 2022) for interactive mapping, and the `spatstat` package (Baddeley et al. 2015) for spatial point pattern analysis.

Chapter 6

Results

6.1 Comparation of frequency and level of danger

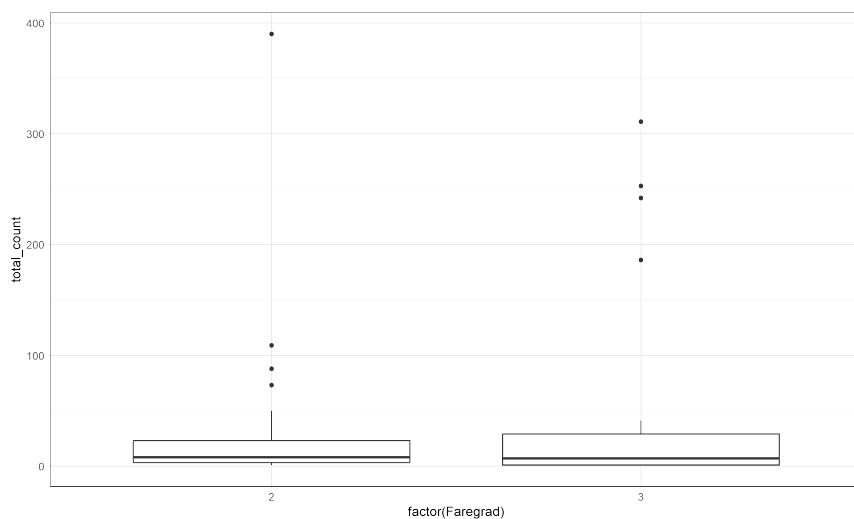


Figure 6.1: Boxplot of Total Count by Avalanche Danger Level

The boxplot in Figure 6.1 shows no notably higher median total avalanche counts for danger level 3 than level 2. Both levels feature outliers, with level 2 having one extreme value and level 3 displaying several higher-valued outliers. However, based on this data, we cannot conclusively state that the danger levels significantly impact the total count of daily avalanches.

Hypothesis test

The result of the ANOVA test is presented in Table 6.1. Based on the p-value of 0.512, we conclude that the avalanche danger level has no significant effect on the number of avalanches. This means that the differences in the means in Table 6.2 of

Table 6.1: ANOVA Table

Source	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Avalanche danger level	1	1914	1914	0.434	0.512
Residuals	88	388237	4412		

Table 6.2: Mean of total observed avalanches for the danger of levels

Danger level	Mean
2	20.3
3	31.9

the danger levels could be due to random chance. We can not reject $H_0 : \mu_2 = \mu_3$.

6.2 A comparative analysis of forecasted terrain aspects and observed avalanche locations

Table 6.3 summarizes the statistical measures for the ratio A over the period of 90 days:

Statistic	Mean	Median	Min	Max	Standard Deviation
Value	0.72	0.8	0	1	0.31

Table 6.3: Statistical summary of the ratio A

The values suggest that, on average, the ratio A is around 0.72, with a median of 0.8, indicating that more than half of the days had a ratio of 0.8 or higher. The minimum value of 0 implies that there were days with no correct predictions, while the maximum value of 1 indicates that there were also days with perfect predictions. The standard deviation of 0.31 reflects the variability of the ratio A , suggesting that there was a moderate spread in the daily accuracy ratios.

Table 6.4 presents the ratios of aspect predictions in all compass directions over the assessed period:

Direction	E	N	NE	NW	S	SE	SW	W
Ratio warned	0.83	0.81	0.91	0.59	0.53	0.72	0.51	0.43
Estimated in PPM	0	0.5	0.44	0.14	0.84	0.33	-0.04	-0.42

Table 6.4: Ratios of correct predictions for each compass direction

From Table 6.4, we can observe that there is no direct linear trend between the warned direction ratio and the estimate of the directions from Section 6.4.

6.3 Poisson regression models

The results of the three Poisson models introduced in Section 5.3 are displayed in Tables 6.5, 6.6 and 6.7, respectively, for Model 1, 2, and 3. The tables contain the coefficients and corresponding p-values that show their significance.

In Model 1, all the coefficients have low p-values, lower than 0.05, which suggests that each predictor is highly significant. Negative coefficients for $p_{30,t}$ (Snowpack stability=30) and $p_{45,t}$ (Snowpack stability=45) categories imply a reduction in the event count, whereas the reference/baseline category is $p_{20,t}$ (Snowpack stability=20), proving that an increase in stability leads to a decrease in avalanche number. Conversely, positive coefficients for $s_{3,t}$ (Avalanche size=3), $f_{S,t}$ (Frequency class=SPECIFIC), and $f_{W,t}$ (Frequency class=WIDESPREAD) suggest an increase. The model shows that Avalanche size=3 gives a higher log count of avalanches than Avalanche size = 2, and for Frequency class, the order from highest to lowest impact on avalanche count is WIDESPREAD, then SPECIFIC, and then ISOLATED. The Intercept provides a baseline for the logarithm of the expected count when other variables are at their baseline category.

Table 6.5: Coefficient estimates from Poisson regression model 1

Coefficient	Estimate	Pr(> z)
Intercept	2.82026	< 2e-16 ***
$p_{30,t}$	-1.00666	< 2e-16 ***
$p_{45,t}$	-1.76064	< 2e-16 ***
$s_{3,t}$	0.85137	< 2e-16 ***
$f_{S,t}$	0.62407	8.71e-05 ***
$f_{W,t}$	1.81764	< 2e-16 ***

Table 6.6: Coefficient estimates from Poisson regression model 2

Coefficient	Estimate	Pr(> z)
β_0	2.98253	< 2e-16 ***
$p_{30,t}$	-1.34937	< 2e-16 ***
$p_{45,t}$	-16.17016	0.925
$s_{3,t}$	-0.89094	1.12e-06 ***
$f_{S,t}$	0.68100	1.86e-05 ***
$f_{W,t}$	1.88505	< 2e-16 ***
$p_{30,t} : s_{3,t}$	1.99085	< 2e-16 ***
$p_{45,t} : s_{3,t}$	16.46563	0.924

Most of the main coefficients in Models 2 and 3 show high significance. The combination of Avalanche size = 3 and Snowpack stability = 30 is significant in Models 2 and 3. The combination of Frequency class = 30 and Avalanche size = 3 is also significant in Model 3. As we move from Model 1 to Model 3, the complexity

Table 6.7: Coefficient estimates from Poisson regression model 3

Coefficient	Estimate	Pr(> z)
Intercept	3.52355	< 2e-16 ***
$p_{30,t}$	-1.38348	< 2e-16 ***
$p_{45,t}$	-17.33264	0.951448
$s_{3,t}$	-1.82389	7.64e-07 ***
$f_{S,t}$	-0.01906	0.938651
$f_{W,t}$	1.50651	9.44e-10 ***
$s_{3,t} : f_{S,t}$	1.09198	0.000718 ***
$s_{3,t} : f_{W,t}$	-16.29400	0.954354
$p_{30,t} : s_{3,t}$	2.16873	< 2e-16 ***
$p_{45,t} : s_{3,t}$	17.62811	0.950621

of the model increases. We compare the AIC values in Table 6.8 to choose the best model. Model 3 has the lowest AIC value. By choosing Model 3 We draw the Poisson distribution graph for three random days from 01-01-2020 to 30-03-2020 shown in Figure 6.2, 6.3 and 6.4. The distribution graphs exhibit a reasonable degree of fit in a certain degree. However, it is evident from the graphs that the model's performance diminishes with higher counts.

Table 6.8: AIC for the three models

Model	AIC Value
Model 1	6449.2
Model 2	6224.9
Model 3	5971.4

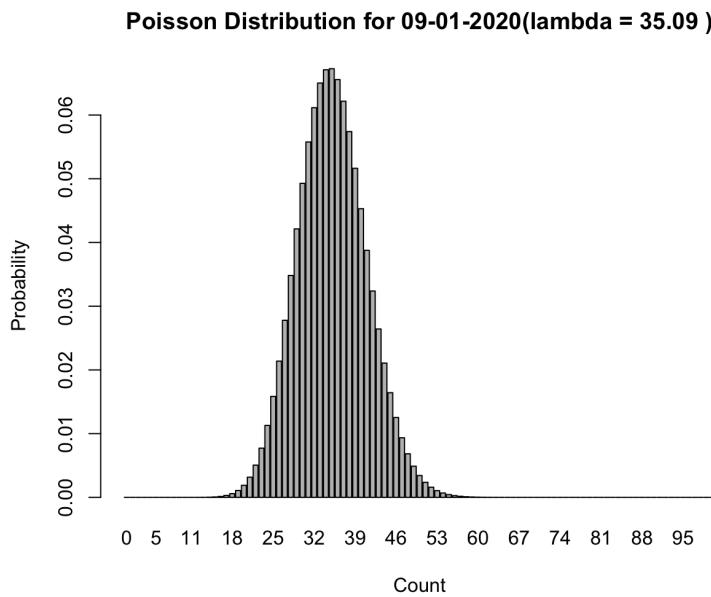


Figure 6.2: Modeled Poisson distribution for day nine ($t=9$). The predicted values were: Frequency class = SPECIFIC, Avalanche size = 3, and Snowpack stability = 30. The observed number of avalanches was 20.

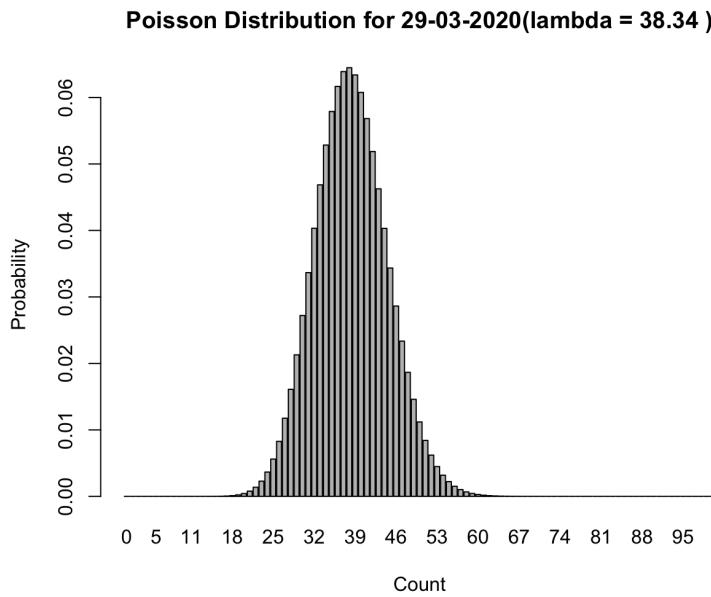


Figure 6.3: Modeled Poisson distribution for day fifty-nine ($t=59$). The predicted values were Frequency class = WIDESPREAD, Avalanche size = 3, and Snowpack stability = 30. The observed number of avalanches was 311.

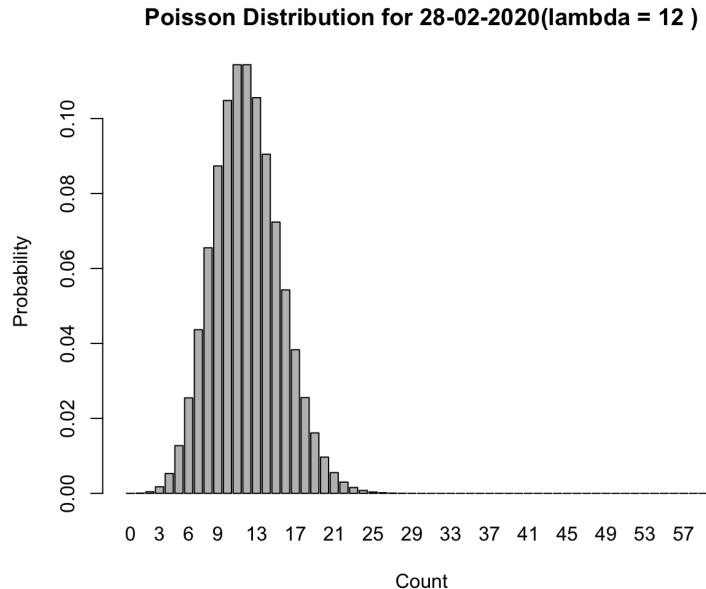


Figure 6.4: Modeled Poisson distribution for day eighty-nine ($t=89$). The predicted values were Frequency class = ISOLATED, Avalanche size = 3, and Snowpack stability = 30. The observed number of avalanches was 0.

6.4 Point process model

The result from the Point process model described in Section 6.4 is displayed in Table 6.9. Based on the result, the estimate for the intercept is -15.21 and is significantly negative, suggesting the baseline level of log intensity ($\ln(\lambda)$) is low when all other variables are zero. On an aspect of terrain East, slope= 0 and elevation = 0 (sea level), the intensity of avalanches is $\exp(-15.2105) = 2.4784 \times 10^{-7}$ per square unit (meter). The intercept has a high absolute Z-value, indicating it is highly significant.

Aspect variables represent the different terrain orientations. The reference category is East (aspect E), which does not appear in the model and will have zero estimate value. F (flat), N, NE, NW, S, SE, SW, and W each have their coefficients. The most notable are N (0.51) and S (0.84), which have high positive coefficients and significant Z-values, suggesting north and south orientations are positively associated with higher log intensity. F and W have significant negative coefficients, indicating these orientations are associated with a decrease in log intensity. The coefficient for F is the lowest, proving that flat areas are associated with low-intensity levels in the model.

Slope Categories represent different ranges of slope steepness. All slope categories from slope at 13° to slope 45° show positive and highly significant coefficients,

Table 6.9: Coefficient estimates from Poisson regression model

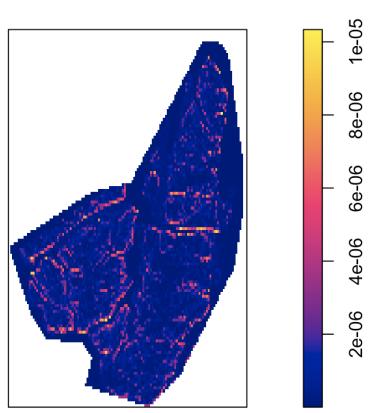
Covariate	Estimate	S.E.	Z-test	Zval
Intercept	-15.2105	0.1213	***	-125.39
A_F	-4.3242	0.7174	***	-6.027
A_N	0.5146	0.0867	***	5.936
A_{NE}	0.4362	0.0867	***	5.032
A_{NW}	0.1421	0.0883		1.609
A_S	0.8399	0.0792	***	10.608
A_{SE}	0.3380	0.0817	***	4.140
A_{SW}	-0.0459	0.0963		-0.477
A_W	-0.4207	0.0975	***	-4.314
$S_{(13-19)}$	1.9720	0.1217	***	16.209
$S_{(19-25)}$	2.8855	0.1161	***	24.860
$S_{(25-31)}$	3.2415	0.1143	***	28.365
$S_{(31-37)}$	2.9710	0.1218	***	24.394
$S_{(37-45)}$	1.9234	0.1619	***	11.878
$S_{(7-13)}$	0.9212	0.1351	***	6.817
$S_{>45}$	0.6006	0.3356		1.790
$E(\vec{x})$	-1.3004	0.0449	***	-28.969

implying steeper slopes are associated with higher log intensity. The highest coefficient is for slope(25-31) $^{\circ}$, indicating this slope range has the strongest association.

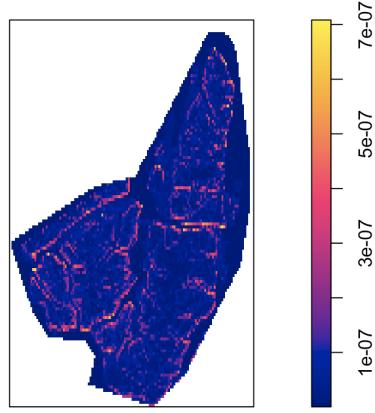
Elevation (E) is scaled by 500 meters in the model. Elevation has a negative coefficient, suggesting for a 500m increase in elevation, the log intensity decreases by 1.3 per square unit area. Zval indicates a significant effect of elevation.

Statistical Significance: The asterisks ('***) denote a high level of statistical significance. Variables without asterisks, like NW and SW, are not statistically significant at the same level. The standard errors are relatively small compared to the estimates, contributing to the high significance levels (as seen in the Z-values).

We can use the fitted Poisson Point Process Model to predict the points in the Lyngen area for 90 days of winter (01-01 to 30-03). The plot of predicted intensity from the model in Equation 5.9 is shown in Figure 6.5a and the corresponding Standard error plot in Figure 6.5b.



(a) Intensity



(b) Standard error

Figure 6.5: Predicted intensity from PPM and the standard error.

Chapter 7

Discussion and Conclusion

In this project, we studied an in-depth analysis of avalanche activities using two distinct types of statistical models: the traditional Generalized linear models (GLMs) and Spatial point process models (PPMs). Integrating these two approaches gave us a more comprehensive understanding of avalanche dynamics by studying the fundamental properties of avalanches and how the number of avalanches and forecasted variables are related.

In isolation, the numerical values of danger levels offer limited insight into avalanche frequency. However, our examination of various models indicates that the factors forming the basis of these danger levels, such as snowpack stability, frequency class, and avalanche size, can yield informative predictions about the expected number of avalanches. Moreover, the interactions among these variables also significantly influence the anticipated number of avalanches. The accuracy of the aspect warnings issued by Varsom is somewhat imprecise when compared to the point process model.

The point process model indicates that terrain aspect, slope, and elevation are key factors influencing the spatial intensity of avalanches. It shows that specific orientations and steeper slopes are more likely associated with higher avalanche intensity. Additionally, the model demonstrates that the impact of the aspect is less significant than that of the slope, supporting David McClung and Peter A Schaefer (2006) statement. Furthermore, a negative correlation exists between elevation and avalanche intensity, suggesting that higher elevations typically have lower avalanche intensities, possibly due to lesser snow accumulation on higher mountains.

Focusing on Spatial Process Analysis, we examined the point patterns over the winter of 2020. This approach highlighted the potential for enhancing avalanche prediction models through detailed daily analysis. Our findings lay the groundwork for future research, where the goal will be to develop more accurate daily avalanche prediction models based on this methodology.

Bibliography

- Lied, K. and K. Kristensen (2003). *Snøskred. Håndboken om snøskred*. Nesbru, Norway: Forlaget Vett og Viten, p. 200.
- EAWS (2024). *EAWS: European Avalanche Danger Scale*. <https://www.avalanches.org/standards/avalanche-danger-scale/>. Last access: 18 Jan 2024.
- Kummervold, Per Egil, Eirik Malnes, Markus Eckerstorfer, Ingar M Arntzen and Fillipo Bianchi (2018). ‘Avalanche detection in Sentinel-1 radar images using convolutional neural networks’. In: *Proc. Int. Snow Sci. Workshop*, pp. 377–381.
- Stoffel, Andreas, Roland Meister and Jürg Schweizer (1998). ‘Spatial characteristics of avalanche activity in an Alpine valley—a GIS approach’. In: *Annals of Glaciology* 26, pp. 329–336.
- Toft, Håvard B, Karsten Müller, Jordy Hendrikx, Christian Jaedicke and Yves Bühler (2023). ‘Can big data and random forests improve avalanche runout estimation compared to simple linear regression?’ In: *Cold Regions Science and Technology* 211, p. 103844.
- McClung, David and Peter A Schaefer (2006). *The avalanche handbook*. The Mountaineers Books.
- Engeset, Rune Verpe (2013). ‘National avalanche warning service for Norway—established 2013’. In: *Proceedings ISSW*, pp. 301–310.
- Müller, Karsten, Frank Techel, Christoph Mitterer, Thomas Feistl, Stefano Sofia, Nicolas Roux, Petter Palmgren, Guillem M Bellido and Lorenzo Bertranda (2022). ‘THE EAWS MATRIX, A LOOK-UP TABLE FOR REGIONAL AVALANCHE DANGER LEVEL ASSESSMENT, AND ITS UNDERLYING CONCEPT’. In.
- McClung, D. and P.A. Schaefer (2023). *The Avalanche Handbook, 4th Edition*. USA: Mountaineers Books.
- ESA (2024). *Sentinel-1*. <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1>. Last access: 18 Jan 2024.
- Eckerstorfer, Markus, Eirik Malnes, Hannah Vickers, Karsten Müller, Rune Engeset and Tore Humstad (2018). ‘Operational avalanche activity monitoring using radar satellites: From Norway to worldwide assistance in avalanche forecasting’. In: *International Snow Science Workshop, Innsbruck, Austria*.
- Eckerstorfer, Markus, Hannah Vickers, Eirik Malnes and Jakob Grahn (2019). ‘Near-real time automatic snow avalanche activity monitoring system using Sentinel-1 SAR data in norway’. In: *Remote Sensing* 11.23, p. 2863.

- Fahrmeir, Ludwig, Thomas Kneib, Stefan Lang, Brian Marx, Ludwig Fahrmeir, Thomas Kneib, Stefan Lang and Brian Marx (2013). *Regression models*. Springer.
- Illian, Janine, Antti Penttinen, Helga Stoyan and Dietrich Stoyan (2008). *Statistical analysis and modelling of spatial point patterns*. John Wiley & Sons.
- Ripley, Brian D (1988). *Statistical inference for spatial processes*. Cambridge university press.
- Renner, Ian W, Jane Elith, Adrian Baddeley, William Fithian, Trevor Hastie, Steven J Phillips, Gordana Popovic and David I Warton (2015). ‘Point process models for presence-only analysis’. In: *Methods in Ecology and Evolution* 6.4, pp. 366–379.
- Rossum, G. van (May 1995). *Python tutorial*. Tech. rep. CS-R9526. Amsterdam: Centrum voor Wiskunde en Informatica (CWI).
- R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: <https://www.R-project.org/>.
- Pebesma, Edzer (2018). ‘Simple Features for R: Standardized Support for Spatial Vector Data’. In: *The R Journal* 10.1, pp. 439–446. DOI: 10.32614/RJ-2018-009. URL: <https://doi.org/10.32614/RJ-2018-009>.
- Cheng, Joe, Bhaskar Karambelkar and Yihui Xie (2022). *leaflet: Create Interactive Web Maps with the JavaScript ‘Leaflet’ Library*. R package version 2.1.1. URL: <https://rstudio.github.io/leaflet/>.
- Baddeley, Adrian, Ege Rubak and Rolf Turner (2015). *Spatial Point Patterns: Methodology and Applications with R*. London: Chapman and Hall/CRC Press. ISBN: 9781482210200. URL: <https://www.routledge.com/Spatial-Point-Patterns-Methodology-and-Applications-with-R/Baddeley-Rubak-Turner/p/book/9781482210200/>.