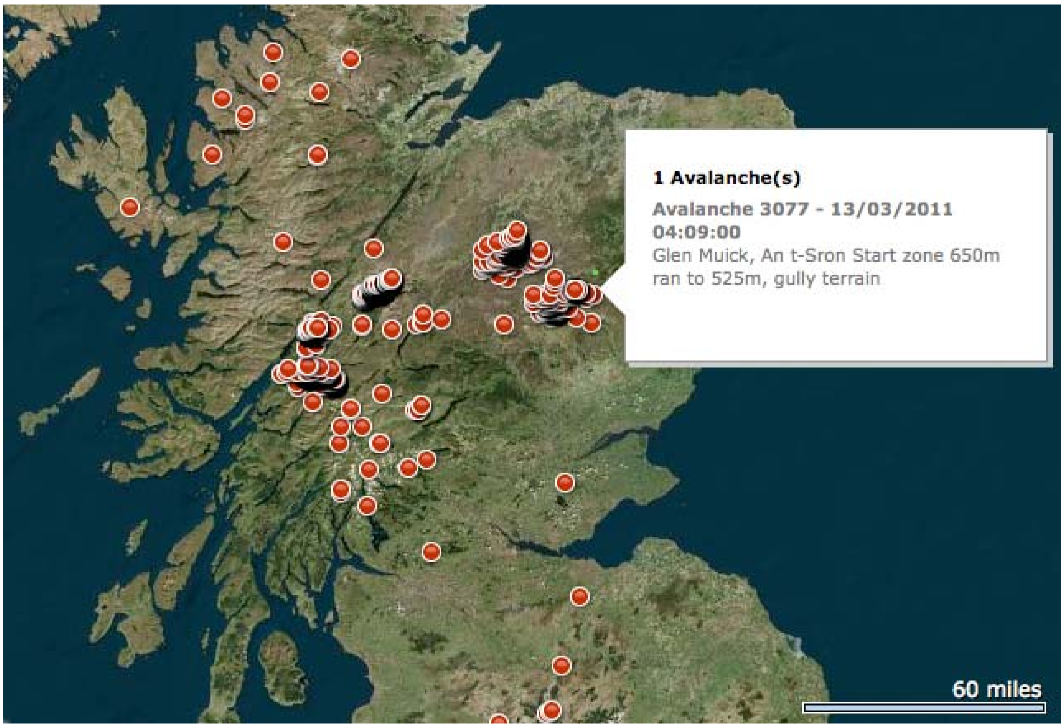
GEO 511: Master Thesis

Investigation of Scottish Avalanche Data Using Spatiotemporal Data Analysis

An Overview of the Current State of Avalanche Activity and Pattern Discovery in Avalanche Data



Eveline Braun
08-704-009

Faculty Representative and Advisor:
Prof. Dr. Ross Purves

Geocomputation, Department of Geography, University of Zurich

27 September 2013

Source of the Figure on the Front Page: SAIS (2013a): Avalanche Map.

PREFACE

Various people have supported me during this thesis in different ways.
I am very thankful to all of you:

- Mark Diggins, Co-ordinator of the SAIS, who made this thesis possible. Thank you for the data, for the encouraging feedbacks and your useful ideas. I hope my work will support you in your future projects as well.
- Prof. Dr. Ross Purves, who made this thesis possible. Thank you for the idea of this interesting topic, for your helpful feedbacks, suggestions and encouragements during the whole time and thanks for the books.
- Richard Essery and Dagan Lev, for the data.
- James Floyer, for the Java code.
- Katharina Jochum and Margret Kraus for proofreading. Without your help this thesis would have been never in acceptable English. Thanks to Monika Bethke-Bühler, Renate Dietzschold and my father as well.
- Emanuela Jochum for the organisation and carrying bunch of papers back and forth.
- Martin Bächtold, for the permission to print this thesis on his printer and my mother for the support.
- The crew and especially Sandra Rota from the Technorama, for the flexibility, consideration, variation and encouragement during the last three years and most notably during the last months and weeks.
- My co-students, for the rewarding lunchs, coffees, teas and great conversations.
- Everyone who had to wait because I said: “I'm sorry; I'll do it in October”.
- My parents, for the enabling of the desired study, the continuous support and sympathy.
- My sister, for her efforts to understand what I am actually doing, for her empathy and handy tips.
- and finally Eugenio, for his infinite patience, steady encouraging, safety and his warm hand.

ABSTRACT

Avalanches are an issue in every Scottish winter. Each season accidents occur, which sometimes end fatal. Therefore, the Scottish Avalanche Information Service (SAIS) was founded in the 80s. The SAIS has been collecting data about avalanches and daily publishes avalanche hazard forecasts for the five popular winter sport areas 'Northern Cairngorms', 'Lochaber', 'Glencoe', 'Southern Cairngorms' and 'Creag Meagaidh'. For forecasts, information of daily snow profiles is combined with meteorological data. Therefore, a data series, lasting over more than 20 years and consisting of different data sets, is available. However, these data have only been used for short time forecasting and never been considered in total. Additionally, the last and only overview of the avalanche activity in Scotland has been published in 1984 (Ward 1984a), as basis for the development of a forecasting model.

Therefore, the first part of this thesis aims to analyse the data set of the reported avalanches in detail and provides an overview of the current avalanche activity. Each attribute is analysed, plotted in R and the characteristics of the Scottish avalanches, including their spatial and temporal distribution, terrain properties, magnitudes and human aspects, are described. Generally, the findings of Ward (1984a) can be confirmed. The majority of the reported avalanches is confined to steeper slopes, higher altitudes and after all to the most popular regions within the SAIS operation area. The majority of the avalanches is relatively small. However, exceptions of much larger avalanches or avalanches which released at far lower gradients and outside of the SAIS areas exist, as well. Additionally, a tendency is found in the data of the last years that the spatial extent of the reported avalanches has increased markedly.

A first look at the snow profile data has shown that various data sources are available, including different attributes and temporal coverages. Therefore, the aim of the second part of this thesis consists of the structuring of these data sets, analysing their quality and combining them into one comprehensive data set that is as complete and consistent as possible. It is looked at the data sets, the meaning of the attributes is clarified and contents are compared. The data are combined, with detailed metadata described and made available for the SAIS for further studies. The data quality is considerably increased in comparison to the original data files. Typing errors, inconsistencies within as well as between attributes are removed and the completeness is considerably extended. However, in the very early years of the SAIS operation as well as between 2003 and 2007, some holes are still present.

To increase the awareness of the existence and hazard related to avalanches in Scotland, the SAIS would like to publish simple, easy understandable and recognisable patterns which describe most hazardous conditions. In other regions, for example in Tirol or Switzerland, such patterns have already been developed, but considering the special maritime Scottish climate with its strong variability, they cannot be applied directly to Scottish conditions. Therefore, the aim of the third part is to discover patterns in the available data which describe most hazardous situations. A Cluster Analysis (CA) is applied on the resulting data set of part B and days with high avalanche activity are grouped, considering weather factors such as air temperature, wind speed and the amount of new snow as well as the snow cover stability as input variables. The comparison of solutions for each SAIS area and the total amount of data results in three obvious patterns including thawing conditions, long cold periods and drifting situations in which much new snow falls. Depending on

the area, between 70% and 95% of the total amount of avalanches can be assigned to one of these clusters. One part of the not assigned avalanches is due to missing data issues, the other avalanches are assigned to a pattern which shows less extreme conditions. The days belonging to this pattern should be analysed further, as this pattern introduces uncertainties into the avalanche problematic and makes the avalanche activity less predictable. However, further factors, more complex processes or temporal sequences, which could not be considered in the Cluster Analysis, might play an important role.

To sum up, this work includes a broad data analysis and describes a variety of aspects concerning the Scottish avalanche activity at the most current state. The data sets are being structured and can be used for future studies without the need for further comprehensive and time consuming data pre-processing steps. Additionally, patterns for hazardous avalanche situations are found, which can support the SAIS to increase the awareness of the avalanche problematic by the public. To enable a profound support to the SAIS, recommendations for improvements are proposed, single steps and results are documented very detailed and additional material is provided.

The aims of the thesis for each of the three parts could be achieved and a new step in the history of research on Scottish avalanches is reached and provides a basis for further projects.

ZUSAMMENFASSUNG

Lawinen sind in jedem schottischen Winter ein Thema. Saison für Saison ereignen sich Unfälle, die manchmal auch tödlich enden. Deshalb gibt es seit Ende der 80er Jahre einen Lawinenwarndienst (SAIS, Scottish Avalanche Information Service), der Informationen über niedergegangene Lawinen sammelt und für die fünf populären Wintersportregionen 'Northern Cairngorms', 'Lochaber', 'Glencoe', 'Southern Cairngorms' und 'Creag Meagaidh' tägliche Vorhersagen der Lawinengefahr publiziert. Informationen von Schneeprofilen werden dazu mit aktuellen Wetterdaten kombiniert. Es existiert somit eine mehr als 20-jährige Datenreihe, welche jedoch einzig für diese kurzfristigen Vorhersagen verwendet und noch nie als Gesamtes betrachtet wurde. Die letzte und einzige Übersicht über die Lawinenaktivität in Schottland wurde 1984 von R. Ward (Ward 1984a) als Basis für die Entwicklung eines Vorhersagemodells erstellt.

Deshalb sollen im ersten Teil dieser Arbeit die Lawinendaten analysiert und eine Übersicht der Lawinenaktivität auf dem aktuellst möglichen Stand gegeben werden. Jedes Attribut aus dem Datensatz der Lawinen wird detailliert angeschaut, mit R dargestellt und die Eigenschaften der Lawinen werden charakterisiert. Die räumliche und zeitliche Verteilung, das Gelände, Grössen und Dimensionen der Lawinen sowie menschliche Aspekte werden beschrieben. Die Resultate von Ward (1984a) können mehrheitlich bestätigt werden. Die Lage der meisten berichteten Lawinen ist beschränkt auf die steileren Hänge, höhere Lagen und vor allem auf die populärsten Regionen innerhalb der SAIS-Gebiete. Die Mehrheit der Lawinen sind relativ klein. Aber Ausnahmen mit grossen Lawinen, oder solchen, die sich auf viel flacheren Hängen lösen und sich ausserhalb der SAIS-Regionen ereignen, sind in einzelnen Fällen ebenfalls möglich. Zusätzlich wird eine Tendenz in den Daten der letzten Jahre ersichtlich, dass sich insbesondere die räumliche Ausdehnung der beim SAIS gemeldeten Lawinen vergrössert.

Erste Blicke auf die Schneeprofildaten haben gezeigt, dass zwar viele verschiedene Datenquellen vorhanden sind, diese jedoch unterschiedliche Attribute beinhalten und verschiedene Zeitspannen abdecken. Deshalb besteht das Ziel des zweiten Teils dieser Arbeit darin, die Datensätze zu ordnen, ihre Qualität zu bestimmen und in einem einzigen Datensatz zusammenzufügen, der möglichst vollständig, einheitlich und von guter Qualität ist. Die Datensätze werden angeschaut, die Bedeutung der Attribute eruiert und Werte verglichen. Die Daten werden dann kombiniert und mit detaillierten Metadaten versehen. Es resultiert eine Datei, welche so viele Informationen wie möglich enthält und dem SAIS für weitere Verwendungen zur Verfügung steht. Die Datenqualität kann im Vergleich zu den originalen Datensätzen erheblich verbessert werden. Tippfehler, Widersprüche innerhalb einzelner wie auch zwischen verschiedenen Attributen werden beseitigt und die Vollständigkeit kann wesentlich erhöht werden. Trotzdem bleiben vor allem in den ersten Jahren der SAIS-Geschichte sowie in den Jahren zwischen 2003 und 2007 Lücken in den Daten vorhanden.

Um das Bewusstsein der Existenz und Gefahr von Lawinen in Schottland bei den Wintersportlern zu stärken, möchte der SAIS einfache, verständliche und gut erkennbare Muster publizieren, welche die gefährlichsten Situationen beschreiben. In anderen Gebieten, beispielsweise dem Tirol oder der Schweiz, wurden solche Gefahrenmuster bereits erarbeitet, können aber wegen dem speziellen maritimen Klima und den sehr stark variablen Bedingungen nicht eins zu eins für Schottland übernommen werden. Im dritten Teil dieser Arbeit wird deshalb eine Clusteranalyse auf die Daten angewendet. Wetterinformationen wie Lufttemperatur, Windgeschwindigkeit oder Neuschneemenge sowie

die Stabilität der Schneedecke dienen als Inputvariabeln, womit die Tage mit hoher Lawinenaktivität in verschiedene Gruppen eingeteilt werden. Der Vergleich der Resultate aus verschiedenen SAIS Regionen sowie den totalen Daten ergibt drei sehr auffällige Muster, welche warme Situationen, lange kalte Perioden und Zeiten mit viel Neuschnee und starker Windverfrachtung beinhalten. Je nach Daten, können zwischen 70% und 95% aller Lawinen einem dieser Muster zugeordnet werden. Die restlichen Lawinen können entweder wegen fehlender Daten nicht zugeteilt werden oder gehören zu einem Muster, das keine solch extremen Bedingungen zeigt. Tage, welche diesem Muster zugeordnet sind, sollten noch weiter analysiert werden. Es ist möglich, dass dieses Muster die Unsicherheit und Unberechenbarkeit aufzeigt, welche Lawinen in sich haben, es kann aber auch sein, dass weitere Faktoren, komplexere Prozesse oder zeitliche Abfolgen, welche in der Clusteranalyse nicht berücksichtigt werden konnten, zusätzlich eine wichtige Rolle spielen und als weitere Muster berücksichtigt werden sollten.

Zusammenfassend beinhaltet diese Arbeit eine breite Datenanalyse und beschreibt verschiedene Aspekte der schottischen Lawinenaktivität auf dem aktuellst möglichen Stand. Es wurde Ordnung in die Vielfalt der Datensätze gebracht, welche nun für weitere Anwendungen zur Verfügung stehen, ohne dass weitere grundlegende, umfassende und aufwändige Datenvorverarbeitungen notwendig sind. Weiter konnten Muster für gefährliche Lawinensituationen gefunden werden, die dem SAIS helfen, das Bewusstsein der Lawinenproblematik in der Bevölkerung zu stärken. Um dem SAIS eine möglichst umfassende Hilfestellung zu bieten, wurden Empfehlungen erarbeitet und die einzelnen Darstellungen, Schritte und Resultate mit grossem Detailgehalt dokumentiert.

Die Ziele aller drei Teile dieser Arbeit wurden erreicht und ein neuer Schritt in der Forschungsgeschichte der schottische Lawinen wurde erarbeitet und kann als Basis für weitere Projekte dienen.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AWS Authomatic Weather Station

CA Cluster Analysis

DEM Digital Elevation Model

EDA Exploratory Data Analysis

ESDA Spatial Exploratory Data Analysis

GIS Geographical Information System

ICSI International Commission on Snow and Ice

KDD Knowledge Discovery in Databases

PCA Principal Component Analysis

SAIS Scottish Avalanche Information Service

SDA Spatial Data Analysis

1 INTRODUCTION

1.1 Motivation

Snow avalanches are known worldwide and are a frequently discussed topic involving many people. Some are directly affected by avalanches, either because they live in a high hazard zone or because they go into avalanche terrain to exercise. Others are engaged in research where avalanches have been on the agenda for a long time. The aim has been to describe the properties of avalanches, explain processes and more recently increasing focus has been put on risk management and forecasting. Therefore current research takes place in different sub fields, and open questions still remain.

Avalanches are an issue in many countries, including Scotland. Scottish winter sport areas are well-known for a variety of activities. Especially winter and mixed climbing has grown into a popular sport, but also hill hikers and ski tourers are making use of the unique atmosphere of a Scottish winter day. Therefore many recreationists are out in the Scottish Highlands and every winter avalanche accidents occur, sometimes ending fatal. As these trends have intensified, a forecast service has been operating for more than twenty years. It publishes daily forecasts for the most popular Scottish winter sport areas.

R. Ward, E. Langmuir and B. Beattie belonged to the first people concerned with avalanches in Scotland. Ward (1980) presents an overview of the avalanche activity in the Cairngorm Mountains at that time and emphasises the need of a forecasting model to "help hill-users to make their own judgements in an informed way" (Ward 1980, 40). Four years later, Ward published a survey of the state of avalanche activity for all of Scotland. He mentions that three different opinions about avalanches were prevalent by the Scots (Ward 1984a, 92):

- "Scottish avalanches (if they exist at all) are small and trivial."
- "Scottish avalanches may be dangerous, but they are predictable since they occur in gullies during thaws [...]".
- "Scottish avalanches are dangerous. They are predominantly of the windslab type and therefore most likely on lee slopes after a blizzard."

The third opinion was the least popular one and nobody had an idea which forms of avalanches are possible in Scotland, how big they are, how frequently and where they release as well as under which conditions they occur (Ward 1984a).

The same year, Ward combined avalanche data with meteorological conditions as basis for the development of a forecast model (Ward 1984b) and in 1988, the Scottish Avalanche Project was founded. All of the people involved agreed on the fact that in Scotland very peculiar conditions are present, which cannot be compared directly to other countries or regions, for example to the Alps.

Scotland is characterised by a temperate, maritime climate, altitudes range from sea level up to sub arctic conditions above 1000 masl and the weather is highly variable and characterised by especially high and frequent wind speeds. Therefore it had been necessary to collect new Scottish data and regular observations near avalanche starting zones began (Barton & Wright 2000). Before the season 1996/1997, the project was absorbed by the Scottish Sports Council and renamed as "Scottish Avalanche Information Service"[1] (Barton & Wright 2000). Today, a more than twenty years lasting data series, consisting of several different data sets, is available. Avalanches are recorded, snow profiles are frequently dug and combined with weather data and daily forecasts are published (SAIS 2013b).

With the growing popularity of Scottish winter sports, the interest in Scottish avalanches and the importance of the work of the SAIS has increased. Starting more than twenty years ago with the publication of forecasts on bulletin boards and in local newspapers for two areas (Barton & Wright 2000), the spatial and temporal coverage has enlarged markedly. Currently, the forecasts are daily updated and extensive publications for the five most popular Scottish winter sport areas 'Northern Cairngorms', 'Lochaber', 'Glencoe', 'Southern Cairngorms' and 'Creag Meagaidh' are provided on the webpage of the SAIS (SAIS 2013b).

Nevertheless, the awareness of the existence of Scottish avalanches is still not present in everybody's mind. The SAIS has a challenging role in providing forecasts, but also in publishing their work with the aim "to make the service known to as many agencies and members of the public as possible. The more the public are made aware of the service, the more it may be utilised, inform and educate" (Diggins 2010, 5). An enhanced awareness of the avalanche hazard offers "the opportunity to plan our excursions and climbs with up to date information" and "enables us to enjoy the Highlands of Scotland with more confidence and safety" (Diggins 2011b, 10).

Target people are mainly reached by publications on the webpage. Some additional background information about avalanches is presented, but it lacks simple, general rules which stick in people's mind. Therefore the SAIS would like to publish a series of patterns showing typical, dangerous situations in which high avalanche hazard is present (pers. comm.: Mark Diggins, Co-ordinator of the SAIS). Working with patterns is not a new approach, as for example the Tiroler Lawinenwarndienst and the SLF made use of it (lawinenwarndienst tirol 2013, SLF 2013) and the value of patterns is explained by Harvey *et al.* (2012, 69) the following: "every day we recognise people we have already met before. The recognition goes without thinking and is very fast and mostly correct. [...] This enormous strength of our brain when interpreting and recognising properties, we can also use for the judgement about avalanche danger."

The work of the SAIS is an important part of the Scottish winter, the interest is growing and a lot of data has been acquired during the last years. However, the data sets have only been used for short-term forecasting and since the survey of the avalanche activity from Ward in 1984, no further comprehensive analysis has been done including all of these different data sets.

Since 1984, the data basis has broadened, the amount of available information has increased enormously and also the knowledge about avalanches in general, from their physical properties to forecasting, has developed significantly. New methods applied in avalanche research have evolved and a variety of research projects in other countries took place. This has shown that there is great potential in the analysis of existing data, in summarising characteristics, looking at distributions or searching for patterns.

Also in Scotland, such a data basis for a thorough analysis is provided by the existing data series. Therefore in this thesis a new approach is applied to extract as much information as possible from the existing data.

[1] The 'Scottish Avalanche Information Service' is commonly abbreviated as SAIS. This official shortening will be used in this work instead of the whole name.

1.2 Aims and Research Questions

The thesis is divided into three parts in which the respective aims are:

A	- to update Ward's study from 1984 providing an overview of the current state of avalanche activity.

B	- to organise and sort the variety of available data sets. - to assess the quality of the data. - to prepare one comprehensive, well described data file in which as much information as available is included, which is updated, consistent and as complete as possible.

C	- to reveal patterns in the data which describe most hazardous conditions and help the SAIS in communicating with the public.

For each of these parts specific research questions are formulated:

A	- What can be said about general avalanche activity in Scotland? - Which characteristics do the Scottish avalanches show?

B	- Which properties do the different data sets show? - Which limitations does the data have? - Is it possible to combine the different data sets into one comprehensive data set? - Is it possible to ensure satisfying data quality in the resulting data set?

C	- Which patterns does the data show? - Are there typical factors or combinations of factors concerning topography, weather, snow cover and humans which lead to an increased avalanche activity?

Different methods are applied to answer these questions, reaching from descriptive statistics and visualisation techniques over spatial data analysis, multivariate analysis methods such as Principal Component Analysis (PCA) and Cluster Analysis (CA) to qualitative techniques. The data analysis is clearly structured, but the approach is explorative and iterative.

1.3 Structure

In a first step the theoretical framework for the topics crucial to this thesis is proposed in chapter 2. Some facts about avalanches, avalanche data and its analysis are described. A summary of the current state of research is provided and relevant studies with their main methods and results are presented. Chapter 3 provides more information about the SAIS and the study areas whereas the available data sets are described in chapter 4. Chapter 5 includes a synthesis in which the different issues presented in chapters 2, 3 and 4 are brought together in order to describe the research gap which this thesis aims to fill.

Methods are presented in chapter 6 and the results are summarised in chapter 7. The thesis ends with the discussion given in chapter 8, where interpretations about data, methods and results are provided and comparisons with other studies are given. In the very last part, chapter 9, the most important points of this thesis are summarised, recommendations for the SAIS and suggestions for future research are provided.

2

THEORY

To answer the research questions formed in section 1.2, background knowledge of different topics has to be put together. Basics about avalanches, their characteristics, building factors, related processes and classification as well as a methodical background for analysing avalanche data are of interest. As avalanches are a spatial phenomenon, spatial characteristics of avalanche data have to be considered as well.

In this chapter, characteristics of avalanches most important in the context of this thesis are summarised. Classification parameters commonly used to describe characteristics of avalanches are presented and factors which influence avalanche activity are described. Other studies where avalanche data has been analysed are presented (see table 2.1 for an overview) and the current state of research is shown. When ever possible, references and specialities concerning Scottish conditions are being provided.

2.1 Avalanches

Avalanches belong to natural hazards and can be defined as "falling masses of snow that contain rocks, soil, or ice" (McClung & Schaerer 1993, 11). Avalanche characteristics can vary heavily and the actual release of an avalanche is dependent on various factors and depending on the conditions. Avalanches are relevant to us from the point they are related to human activity, as soon as people or infrastructure are endangered.

2.1.1 Avalanche Properties and Classification

Generally, a distinction is made between ***loose snow*** and ***slab*** avalanches. Loose snow avalanches start with little snow amount at a single point, entrain more and more snow on their way and become bigger while running down the slope. Slab avalanches start due to a failure in the snowpack at a horizontal fracture line and the whole mass of snow runs down the slope, whereby the width stays relatively constant (McClung & Schaerer 1993).

Further, avalanches can be categorised into ***full depth*** or ***surface*** avalanches, depending on the depth of the moving snowpack. The liquid water content of the snowpack defines if an avalanche is ***dry*** or ***wet***. Depending on the form of the avalanche path, avalanches can be ***unconfined*** or ***channelled***. The movement of snow is described either as ***airborne*** or as ***surface*** movement (Barton & Wright 2000). Some of these morphological characteristics are shown in figure 2.1.

The ***size*** of an avalanche can be measured using different variables such as defined size scales, the destructive force, the size relative to the path, widths, fracture depths, volumes, vertical falls, runout distances or deposit dimensions (Greene *et al.* 2010a).

Table 2.1: Overview of relevant literature for the theoretical background.

Author	Year	Title	Subject: Avalanche Background	Subject: Avalanche Factors	Subject: Avalanche Analysis	Subject: Pattern Analysis	Subject: Pattern Examples
Watanabe, S.	1985	Pattern Recognition: Human & Mechanical				✖	
Föhn, P.	1992	Characteristics of Weak Snow Layers or Interfaces		✖ snow			
Mock, C.J. & Kay, P.	1992	Avalanche Climatology of the Western United States, with an Emphasis on Alta, Utah				✖	
McClung, D.M. & Schaerer, P.	1993	The Avalanche Handbook	✖	✖ terrain / weather			
Fredston, J. et al.	1994	The Human Factor-Lessons for Avalanche Education		✖ human			
Stoffel, A. et al.	1998	Spatial Characteristics of Avalanche Activity in an Alpine Valley - A GIS Approach			✖		
Barton, B. & Wright, B.	2000	A Chance in a Million? Scottish Avalanches	✖	✖ terrain / weather			
Jain, A. et al.	2000	Statistical Pattern Recognition: A Review				✖	
Kleemayr, K. et al.	2000	Lawinenprognose mit statistischen und selbstlernenden Verfahren im Projekt NAFT					✖
Mock, C.J. & Birkeland, K.W.	2000	Snow Avalanche Climatology of the Western United States Mountain Ranges				✖	
Schweizer, J. & Lütschg, M.	2000	Measurements of Human-Triggered Avalanches from the Swiss Alps		✖ human			
Birkeland, K.W. et al.	2001	Avalanche Extremes and Atmospheric Circulation Patterns				✖	
Schweizer, J. & Lütschg, M.	2001	Characteristics of Human-Triggered Avalanches		✖ human			
Signorell, C.	2001	Skifahrerlawinenunfälle in den Schweizer Alpen - Eine Auswertung der letzten 30 Jahre				✖	
Wiesinger, T. & Schweizer, J.	2001	Snow Profile Interpretation		✖ snow			
Esteban, P. et al.	2002	Application of Multivariate Statistical Techniques for the Characterisation of the Surface Circulation Patterns Related to Heavy Snowfalls in Andorra, Pyrenees				✖	
Johnson, R. & Birkeland, K.	2002	Integrating Shear Quality into Stability Test Results		✖ snow			
Laternser, M. & Schneebeli, M.	2002	Temporal Trend and Spatial Distribution of Avalanche Activity During the Last 50 Years in Switzerland			✖		
Maggioni, M. & Gruber, U.	2002	The Influence of Topographic Parameters on Avalanche Release Dimension and Frequency		✖ terrain			
McCammon, I.	2002	Evidence of Heuristic Traps in Recreational Avalanche Accidents		✖ human			
McClung, D.M.	2002	The Elements of Applied Avalanche Forecasting, Part I: The Human Issues		✖ human			
McClung, D.M.	2002	The Elements of Applied Avalanche Forecasting, Part II: The Physical Issues and the Rules of Applied Avalanche Forecasting		✖ human			
Hägeli, P. & McClung, D.M.	2003	Avalanche Characteristics of a Transitional Snow Climate - Columbia Mountains, British Columbia, Canada			✖		
McClung, D.M.	2003	Magnitude and Frequency of Avalanches in Relation to Terrain and Forest Cover		✖ terrain			
McCollister, C. et al.	2003	Exploring Multi-Scale Spatial Patterns in Historical Avalanche Data, Jackson Hole Mountain Resort, Wyoming			✖		
Schweizer, J. & Jamieson, J.	2003	Snowpack Properties for Snow Profile Interpretation		✖ snow			
McCammon, I.	2004	Heuristic Traps in Recreational Avalanche Accidents: Evidence and Implications		✖ human			
McCollister, C.M.	2004	Geographic Knowledge Discovery Techniques for Exploring Historical Weather and Avalanche Data			✖		
Tase, J.E.	2004	Influences on Backcountry Recreationists' Risk of Exposure to Snow Avalanche Hazards		✖ human			
Adams, L.	2005	A Systems Approach to Human Factors and Expert Decision-Making within Canadian Avalanche Phenomena		✖ human			
Esteban, P. et al.	2005	Atmospheric Circulation Patterns Related to Heavy Snowfall Days in Andorra, Pyrenees				✖	
Oller, P. et al.	2006	The Avalanche Data in the Catalan Pyrenees. 20 Years of Avalanche Mapping			✖		
Schweizer, J. et al.	2006	Spatial Variability - So What		✖ snow			
Eckert, N. et al.	2007	Hierarchical Bayesian Modelling for Spatial Analysis of the Number of Avalanche Occurrences at the Scale of the Township			✖		
Hägeli, P. & McClung, D.M.	2007	Expanding the Snow-Climate Classification with Avalanche-Relevant Information: Initial Description of Avalanche Winter Regimes for Southwestern Canada			✖		
Jomelli, V. et al.	2007	Probabilitstic Analysis of Recent Snow Avalanche Activity and Weather in the French Alps				✖	
McCammon, I. & Hägeli, P.	2007	An Evaluation of Rule-Based Decision Tools for Travel in Avalanche Terrain		✖ human			
Schweizer, J.	2007	Neuer Schneedeckentest - "Nieten" in der Schneedecke suchen		✖ snow			
Garcia, C. et al.	2008	Atmospheric Patterns Leading Major Avalanche Episodes in the Eastern Pyrenees and Estimating Occurrence				✖	
Harvey, S.	2008	Sicherheit im Bergland - Mustererkennung in der Lawinenkunde					✖
Schweizer, J.	2008	Profilinterpretationen		✖ snow			
Fierz, C. et al.	2009	The International Classification for Seasonal Snow on the Ground		✖ snow			
Nairz, P.	2009	Avalanche Patterns as Aid in Avalanche Forecasting					✖
Greene, E. et al.	2010	Snow, Weather, and Avalanches: Observation Guidelines for Avalanche Programs in the United States		✖			
Nairz, P. & Mair, R.	2010	Was sind Gefahrenmuster?					✖
Schweizer, J. & Jamieson, J.	2010	Snowpack Tests for Assessing Snow-Slope Instability		✖ snow			
Casteller, A. et al.	2011	Reconstructing Temporal Patterns of Snow Avalanches at Lago del Desierto, Southern Patagonian Andes			✖		
Guy, Z.M.	2011	The Influence of Terrain Parameters on the Spatial Variability of Potential Avalanche Trigger Locations in Complex Avalanche Terrain		✖ terrain			
Vontobel, I.	2011	Geländeanalysen von Unfalllawinen		✖ terrain			
Guy, Z.M.	2012	Avalanche Trigger Locations in Complex Terrain		✖ terrain			
Guy, Z.M.	2012	Relating Complex Terrain to Potential Avalanche Trigger Locations		✖ terrain			
Harvey, S. et al.	2012	Lawinenkunde - Praxiswissen für Einsteiger und Profis zu Gefahren, Risiken und Strategien	✖				
Harvey, S.	2013	Vier Muster für Lawinengefahr					✖

As avalanches are a spatiotemporal phenomenon, each avalanche is associated to a ***location***, the ***terrain characteristics*** of the avalanching slope as well as the specific ***time of release*** can be described.

Referring to Scotland, Barton & Wright (2000) in Barraclough *et al.* (2010) mention that slab and wet snow avalanches are most common in Scotland and that the majority are full depth avalanches.

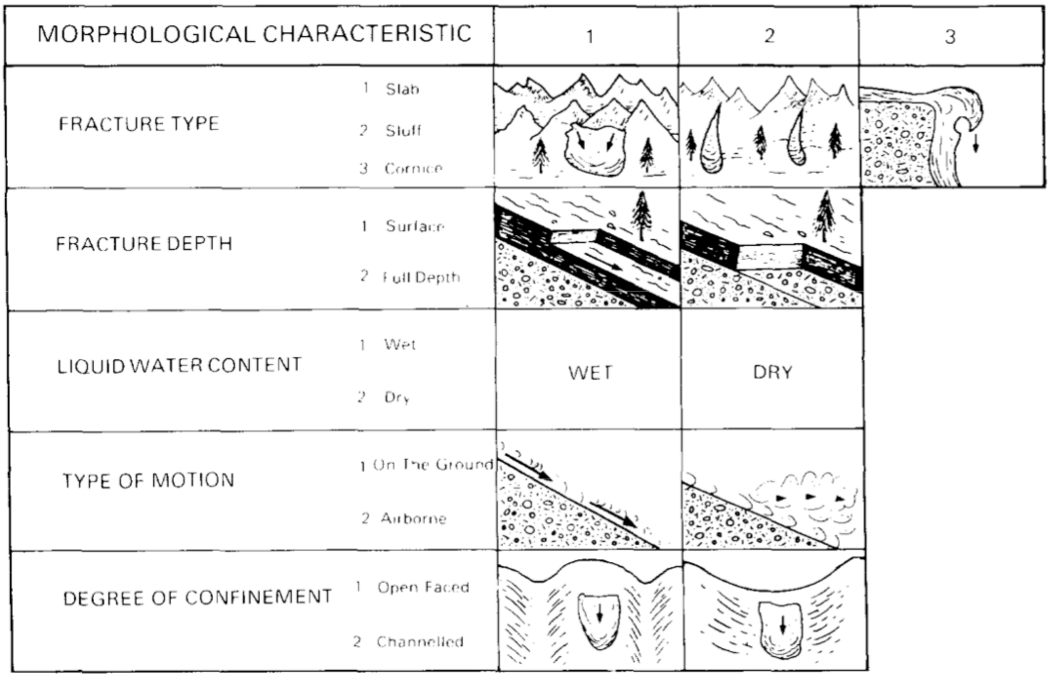


Figure 2.1: Morphological characteristics for describing avalanche properties (Ward 1984a, 98).

2.1.2 Avalanche Factors

When a slope is avalanching, many different factors play together and complex processes take place. For a long time, ***terrain*** and ***weather*** which influence the ***snow cover*** have been seen as most decisive ones. The terrain is a stable factor, whereas weather and snow cover show temporal and spatial variability on different scales (Ward 1984b). More recently, the ***human*** factor has been added and considered to play an important role as well, not least because many avalanches are triggered by humans (e.g. Schweizer & Lütschg 2000). Different aspects of these four factors relevant to this thesis are summarised in the following sections.

2.1.2.1 Terrain

Characteristics of a slope are defined by the terrain. The terrain varies heavily and no two slopes are identical. The following factors define slope properties and can play an important role in determining avalanche activity:

Gradient influences two different factors relevant to avalanche activity and constrain the range of slope angles to the ones on which enough gravity force is and sufficient snow is available to build avalanches. The steeper a slope, the faster a snow mass moves but the less snow can be hold over longer periods. However, exact thresholds provided in literature differ slightly (see table 2.2 and figure 2.2).

Referring to Scotland, Barton & Wright (2000, 43) state that avalanches "start most commonly on slopes of an inclination between 30° and 45°, with the maximum occurrence at about 37°" and "there seem to be no upper limits of angles for avalanches in Scottish gullies" (Barton & Wright 2000, 75).

Aspect influences avalanche activity mainly because of the amount of incoming sun energy. Northerly aspects are prone to avalanches as they normally do not have sun for a long time, are shadier, colder, conditions are less variable and instabilities or weak layers persist longer combined to other ones (Harvey *et al.* 2012, McClung & Schaerer 1993). Harvey *et al.* (2012) and McClung & Schaerer (1993) mention south-facing slopes because of the direct exposure to

Table 2.2: Typical gradient values for avalanching slopes from literature.

Author	Minimum Inclination	Mean Inclination	Maximum Inclination
Fyffe (2000)	20° (30° more likely)	38°	60° (45° more likely)
Harvey et al. (2012)	30°	35°	
Maggioni & Gruber (2002)	30° (case study: 20°)	33-38°	60° (case study: 51°)
McClung (2003)		38° / 39° / 41°	
McClung & Schaerer (1993)	25°		60°
Vontobel (2011)		35°	60°

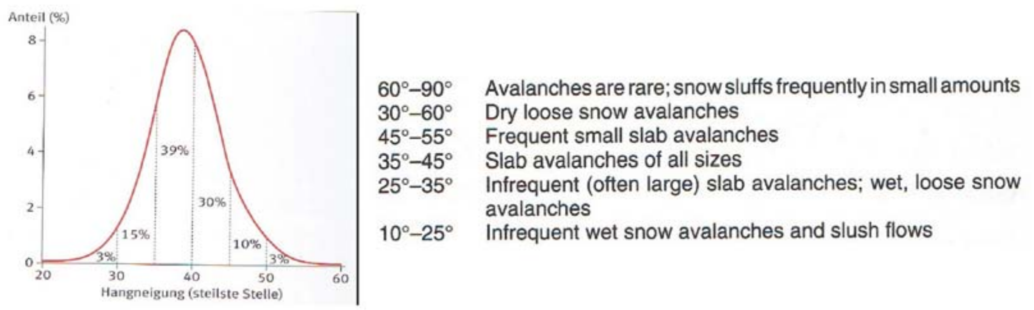


Figure 2.2: Distribution of gradients for possible avalanche release zones. (Left: Harvey et al. *2012, 60 | Right: McClung & Schaerer 1993, 92).*

sunlight for example on the first sunny day after snowfall, lee slopes on which drifted snow is accumulated and lee slopes under cornice roofs as hazardous aspects as well.

Therefore avalanches can generally release at every aspect, but tendencies for more hazardous ones exist. These strongly depend on the local situation and the distribution of aspects at specific locations, shown in case studies where results vary highly. Stoffel *et al.* (1998) see potential avalanche starting zones favoured on S and SE facing slopes in the Engadine valley in the Swiss Alps. The study of Maggioni & Gruber (2002) results in the following avalanche prone aspects: N: three times, NE/SW/W/NW: each two times, E/S: each once. Oller *et al.* (2006) found equal proportions at eastern, southern and western aspects and only few avalanche releases on northern oriented slopes which "is logical, taking into account the north-south direction of the main valleys" (Oller *et al.* 2006, 311).

Altitude has an influence on several factors which determine avalanche activity. No strict altitudinal range with high avalanche activity can be defined, as conditions at certain altitudes are highly dependent on the specific location, topography and the local climate. According to McClung & Schaerer (1993, 95) avalanches can occur "on upper slopes when conditions on lower slopes are stable and vice versa". It can be distinguished between three different altitudinal belts: areas below the timberline, big and homogenous slopes above the timberline and summit regions with ridges and saddles at higher elevations. These slopes behave different in relation to avalanching because wind, temperature, amount of snow and variability in the snow cover change. Generally, wind speeds and precipitation tend to be higher at higher altitudes and temperature decreases with increasing altitude. The lower the altitude the higher the possibility that precipitation has the form of rain instead of snow (Harvey *et al.* 2012).

In Scotland, the altitude ranges from 0 masl to 1344 masl. Most flat areas are below 900 masl where a change in topography occurs and the slopes become steeper. Thus, the Highlands on which the vast majority of avalanches occur lie above 900 masl (pers. comm.: Mark Diggins, Co-ordinator of the SAIS).

Roughness of the surface influences the development of a snow cover and its resistance to sliding. According to McClung & Schaerer (1993) open slopes, thinly forested areas, deep snowpacks with buried anchors and smooth slopes without anchors are prone to avalanches. On a slope with big boulders, a continuous snow cover can not develop directly above the ground and the chance for movement is low as long as the boulders are not covered entirely. In contrast, on a meadow with long grass or on rocky slopes, the cohesion between snow and ground is low and it takes only little until a snow mass starts to move (McClung & Schaerer 1993).

Curvature describes the shape of the terrain which can be convex or concave. Vontobel (2011) found that avalanches mostly release on planar slopes or on transitions from concave to convex slopes leading to dangerous locations, for example, below ridges.

Distance from a ridge has been analysed by few authors. The study of Vontobel (2011) shows that many avalanches release close to ridges with the majority having in average less than 200 meters in distance. Maggioni & Gruber (2002) found distances to ridges for potential release areas at values of 20, 60, 100 and 500 meters.

2.1.2.2 Weather

According to McClung & Schaerer (1993, 17) the "primary weather and atmospheric factors contributing to avalanche formation include precipitation patterns and intensity, wind direction and wind speed, sensible heat and radiation heating or cooling on snow". These factors always play together and different combinations can lead to a variety of conditions (Harvey *et al.* 2012). Most important meteorological factors are briefly described:

New Snow is built in the atmosphere, falls down on the ground, accumulates and forms a layer of new snow on the old surface. The two most important influences are the additional load, depending on the amount and density of the new snow, and the quality of bonding to the old surface, depending on the characteristics of the old surface and the new snow crystals. Further, the snow cover is influenced by the conditions during accumulation, such as wind or temperature (Harvey *et al.* 2012).

New snow can take various forms and one possible classification according to the *International Commission on Snow and Ice (ICSI)* distinguishes between ***columns, needles, plates, stellar crystals, irregular particles, graupel, hail*** and ***ice pellets*** (McClung & Schaerer 1993). Four categories of new snow are mentioned in Barton & Wright (2000): ***new snow*** near its initial crystal form which does not have been influenced heavily by wind or temperature changes, ***wet snow*** which is close to the melting point due to rising temperatures, ***broken crystals*** which have been affected by strong winds, and ***rimed crystals*** which have gone through freeze-melt processes.

Several conditions determine the avalanche danger after a new snow event. The avalanche activity increases with the strength and intensity of snow falls, with the density and lower temperature of the new snow, larger temperature difference between old and new snow and with the softness and coarseness of the old surface (Harvey *et al.* 2012). Barton & Wright (2000, 42) mention that in Scotland "almost all large, dry slab-avalanches follow a period of heavy snowfall or considerable drifting."

Harvey *et al.* (2012) define the critical new snow depth for the release of avalanches between 10 cm and 20 cm when otherwise favourable conditions are present, between 20 cm and 30 cm when middle favourable conditions exist and between 30 cm and 50 cm when avalanches are not favoured otherwise.

Rain has a large influence on the snow cover and is able to increase the avalanche danger over a short time as the snow is warmed up quickly and looses its strength rapidly. Additionally, the snow gets much heavier which leads to an increased stress on the snow cover. If rain falls

Table 2.3: Thresholds of wind speeds to enable drifting from literature.

Author	Loose New Snow	Loose Old Snow	Dense Snow
Barton & Wright (2000)	20 mph		
Harvey et al. (2012)	10 mph	25 mph	40 mph
McClung & Schaerer (1993)	12 mph		> 60 mph

on dry snow, it pours down through permeable layers and accumulates above stronger layers resulting in heavy destabilisation. If the snow is already wet, the influence of the additional rain water is not that strong. In most cases, the water then flows down to the ground and leaves the snow cover (Harvey *et al.* 2012).

Due to the maritime climate, rain is an important factor considering avalanche activity in Scotland, as it may influence the snow cover during the whole winter (Purves *et al.* 2003).

Wind is according to Barton & Wright (2000, 36) "probably the most important factor influencing avalanches in Scotland". Two important characteristics of the wind influence the formation of avalanches: wind speed and wind direction. The main process driven by wind is the transport of snow which is called drifting.

Defining thresholds for wind speeds that lead to snow drifting is not an easy task. Generally, "snow is picked up in places where the wind accelerates, and snow is deposited in deceleration zones" (McClung & Schaerer 1993, 25), but only few authors define specific thresholds (see table 2.3[1]). Snow crystals are damaged by turbulence in the air and by rolling or saltation along the ground. As soon as the wind speed is getting lower, drifting snow is accumulated as compact windslab mostly on lee slopes, in corries or hollows. Barton & Wright (2000, 37) state that "a deposition rate of 3 cm of new snow per hour is considered dangerously high", especially because "newly and rapidly deposited windslabs seem to be particularly unstable". Depending on wind speed, windslabs can have different densities and it can be distinguished between hard and soft slab, according to their ability to carry a person or not. After deposition, metamorphism processes are similar to those in new snow layers (Barton & Wright 2000).

Wind direction defines in which direction the snow is transported and thus which slopes are acting as lee zones (McClung & Schaerer 1993).

Summarising, Harvey *et al.* (2012) emphasise that wind is able to transform loose snow to dangerous windslabs in only few hours and to bring either warm or cold air to a certain location in short time which influences the snow surface. Generally, drifting as well as accumulation processes make the snow cover unstable (McClung & Schaerer 1993).

Air Temperature normally has a small impact on the snow cover as it is only one part of the energy balance of a surface. However, Harvey *et al.* (2012) present five situations in which air temperature may have a bigger influence on the snow cover.

- ***Thaws*** until about 0 °C introduced by rain or strong radiation and high air temperature lead to changes in the surface layer of the snow cover. The snow gets warmer and more deformable, leading to higher stress on deeper layers and instability in the snow pack.
- ***Long warm periods*** lead to melting processes in the snow cover and water can destabilise the snow.
- During ***long cold periods*** with snow temperatures below -5 °C to -10 °C, the snow gets brittle and is deformable no longer. Snow settle and stabilisation processes develop slowly and unstable conditions prevail for long time. With even deeper temperatures, the snow cover is influenced by constructive metamorphosis as well (see section 2.1.2.3).
- ***Fast cooling*** or ***fluctuating temperatures*** lead to stabilisation of the snow cover as the bonding between single crystals is strengthened.

[1] Wind speeds provided in the original literature are not always given in [mph]. They are converted from [m/s] or [km/h] into [mph] and rounded to whole numbers for better readability.

Additionally, air temperature may have a higher influence on the snow surface when winds increase the exchanging of heat.

Heat exchange with the atmosphere can take place as conduction, convection or radiation, and influences the stability of the snowpack heavily. ***Radiation*** interacts with the snow surface as direct short-wave radiation or as indirect long-wave radiation and can either lead to the building of weak layers on the snow surface or directly trigger an avalanche by warming or cooling (McClung & Schaerer 1993). Direct radiation is dependent on cloud cover, mist and time of year and day. The main part of the incoming radiation is reflected by the bright snow surface, the rest is absorbed by the snow cover and transformed into heat. The bigger the angle of incoming radiation, the stronger the heating. A thick cloud cover leads to warming in the snow due to limited emission. With a clear sky, the emission is high and the snow surface cools down (Harvey *et al.* 2012).

Conduction processes can be neglected as they are very small. ***Convection*** takes place due to turbulent exchange of sensible heat at the surface or by condensation resulting from diffusion of water vapour. The formation of surface hoar when air with high water vapour content condenses on a cold snow surface is a typical example (McClung & Schaerer 1993).

Most of these weather factors play their role on different scales and are highly variable in space and time due to "meteorological conditions and topography, predominantly radiation and wind, the climate in general at larger scale" (Schweizer *et al.* 2006, 374). In relation to avalanches, the latter can be defined by the snow climate. Generally, three different snow climates are distinguished: the ***continental***, ***maritime*** and ***transitional*** one (McClung & Schaerer 1993). Mock & Birkeland (2000) describe the continental climate as having relatively low temperatures, little snow, and frequent clear skies. This leads to weak layers that can persist over long periods. Maritime climates show higher temperatures, heavier snowfalls and more clouds. Additional characteristics are deep snow covers and rain as well as cold Arctic air which can influence the snow cover throughout the whole winter. Therefore weather changes are fast, "snow covers are often very unstable with rapidly fluctuating instability" and avalanches are frequently the result of heavy snow storms, rainfall into the snow cover or the formation of ice layers (McClung & Schaerer 1993, 18). Transitional climates are located between the continental and maritime climate, but they can lean towards either, depending on the local climate (Mock & Birkeland 2000).

Hägeli & McClung (2003, 255) are convinced that "even though the *avalanche climate* and the *snow climate* of an area are closely related, a clear distinction between these two terms should be made." The avalanche climate should be used to describe the snow climate of an area more thoroughly, including characteristics such as internal structure of the snow pack, weaknesses and avalanche activity. Therefore Hägeli & McClung (2007) suggest the new term ***avalanche winter regime*** as a new classification scheme, which allows to consider also local snowpack structures.

2.1.2.3 Snow Cover

The third important factor influencing the avalanche activity is the snow cover itself. The life of snow starts in the atmosphere and the properties and form of snow crystals are highly dependent on temperature and humidity conditions which occur during their building and transport. McClung & Schaerer (1993, 41) state the following: "changes in crystal types during storms, including changes in the amount of riming, can create conditions where one layer does not bond well to the next. This can be of significance in prediction of snow stability. [...] The integrated effects of crystal form, riming, and breakage can all contribute to instability in new snow and its bonding characteristics with old snow layers".

After deposition, many processes can change the properties of snow as well. Because the snow temperature is close to the melting point, only little variation in temperature, wind, sun energy or rain can lead to considerable changes in the snow cover (Harvey *et al.* 2012) and the snow "is in a continuous state of transformation, known as metamorphism" (Fierz *et al.* 2009, 3).

Generally, it is distinguished between four types of metamorphism. The ***mechanical metamorphism*** describes mainly the effect of wind on snow crystals leading to broken crystals.

Dry snow metamorphism occurs when small temperature gradients are present and leads to the rounding of snow crystals. Through sublimation, water vapour moves from the complex branches of the snow crystals to their centres. In sintering processes, bonds can be built between single crystals, which stabilise the snow cover (Barton & Wright 2000). With temperatures between -5 °C and 0 °C, these processes develop with high velocity (Harvey *et al.* 2012).

If a strong temperature gradient of more than 10 °C / m is present in the snow pack, ***constructive metamorphism*** takes place and kinetic growth of the snow crystals lead to fragile ones such as facets, hoar or cup crystals. These crystals can act as weak layers in the snow pack. Referring to Scotland, Barton & Wright (2000, 34) state that "until recently, it was thought that with the exception of surface hoar, kinetic growth forms did not play a significant part in Scottish avalanches. However, systematic observation, particularly of shallow snowpacks, has shown facets to be widespread and other forms not infrequent".

Dry snow and constructive metamorphism develop without any influence of melting and liquid water, in contrast to the ***melt-freeze metamorphism***. The snow pack is influenced by cycles of melting and freezing, including the effect of liquid water. Snow crystals are then surrounded by water, bondings get weaker and when freezing, this combination leads to a strong structure consisting of large and coarse grains, which in Scotland is often called névé, a synonym for firn. This structure is stable because all grains are bonded or sintered and it is the normal condition of Alpine snow. In Scotland, however, the degree of melting and re-freezing is much stronger and often acts as a dominant process (Barton & Wright 2000). Strongest melting is introduced by warm winds, rain or strong solar radiation in spring time. In Scotland "the mechanism of melting and re-freezing is a most important one for stabilising the snowpack during winter and the 'freeze' phase of the cycle gives our best winter climbing conditions." (Barton & Wright 2000, 36).

The snow pack is structured in different layers which have their own characteristics and either favour avalanches or stabilise the snowpack. There are two special phenomena directly related to the snow cover:

Weak layers or interfaces in the snow pack are a sign of instability and "much of the practical defence of the mountaineer against avalanches rests on attempts to identify the weakest layer in the snowpack" (Barton & Wright 2000, 45). A whole layer can be unstable, because it consists of fragile or unconsolidated crystals which have little bonding and big air spaces between them. Weak layers typically are soft, permeable to light because of big air spaces and therefore prone to fast breaking or collapsing. Interfaces can be unstable as "any major difference of hardness, wetness or crystal size between adjacent layers indicates a poor adhesion between these layers" (Barton & Wright 2000, 44/45).

According to Ward (1980) weaknesses in Scottish snow covers can arise for a number of reasons. Slush layers, cohesionless grains as a result of a new snow layer on surface hoar or the burying of hail into the snow pack are some examples he mentions.

Cornices are a very common phenomenon in Scotland. When wind is blowing over a ridge, loose snow can be accumulated as windslab on lee slopes with wind speeds between 10 mph and 55 mph (~5 m/s to ~25 m/s) (Barton & Wright 2000, McClung & Schaerer 1993). Below the cornice, a scarp slope is built with a gradient of about 52°. Cornices can overhang and have an intrinsically unstable structure. The hazard for cornice collapse increases the more overhanging they are, the steeper the scarp slope and the newer the formation is or when thawing conditions exist (Barton & Wright 2000). In most cases, cornices form on ridge crests but according to McClung & Schaerer (1993, 29), they can develop "at any place where a sharp change in slope angle is found".

2.1.2.4 Human

Traditionally, when working with avalanche issues in research, the three factors terrain, weather and snow cover have been considered. But in recent years, the focus has been broadened and the human factor was added. Factors related to humans are group size, expertise, equipment and behaviour (Harvey *et al.* 2012) and according to Fredston *et al.* (1994, 473), important variables include "attitude, ego, incorrect assumptions, peer pressure, denial, tunnel vision, complacency, money considerations, poor planning, poor communication, the 'sheep syndrome', the 'horse syndrome' and the 'lion syndrome' ". Some of today's authors concentrate on humans because they are convinced that it even is the most important factor. McClung (2002a, 111) for example states: "since most avalanche accidents result from human errors, no description of avalanche forecasting is complete unless the human component is addressed" and Harvey *et al.* (2012) mention that avalanche accidents are in most cases no coincidence. Only about five percent of the avalanche victims get in a spontaneously released avalanche. In the majority of cases, the avalanche is triggered by the victims or their companions (Harvey *et al.* 2012) and they occur "because the victims either underestimate the hazard or overestimate their ability to deal with it" (Fredston *et al.* 1994). Also in Scotland, this fact seems to be trues, as "avalanches involving people are, in 90% of cases, triggered by their victims" (Diggins 2012).

One of the main researchers on this subject is I. McCammon. He is motivated by the fact that "traditional avalanche education places a heavy emphasis on terrain, snowpack and weather factors on one side. While there is no doubt that this knowledge can lead to better decisions, it is disturbing that the victims in this study that were most influenced by heuristic traps, were those with the most avalanche training. The current and growing emphasis on human factors in avalanche education seems wholly appropriate" (McCammon 2002, 6). The second side of his motivation consists of the phenomenon that "people struggle to explain how intelligent people with avalanche training could have seen the hazard, looked straight at it, and behaved as if it wasn't there" (McCammon 2004, 1) and "that even trained victims commonly ignore obvious clues and fail to take simple precautions" (McCammon 2002, 6).

Therefore McCammon (2002) tries to assess which social factors lure people into going to dangerous regions. He analysed approximately seven hundred incidents which have occurred in the USA between 1972 and 2003 and compared the different circumstances in which the avalanches have released. The author is convinced that simple rules of thumb as well as heuristics exist, which influence decisions during recreation. Heuristics are often used by humans without realising that they are actually applied, because fast and with only little effort they lead to realistic results. But in most cases, their basis consists of only few evidence clues. McCammon (2002) analyses the following heuristics:

Familiarity is the tendency to believe that those things we have always done before are good (McCammon 2002) and to act similarly in comparable situations (McCammon 2004).

Acceptance stands for every human's tendency to make decisions in such a way that respect, acceptance and reputation is ensured by others. When gender plays a role, this heuristic can be especially pronounced (McCammon 2004).

Consistency stands for the characteristic of a human to see everything what is planned and aimed for the day as good and right. If a person makes a choice and has to make it a second time later, he or she tends to make the decision in the same way than before (McCammon 2002).

Expert halo describes the fact that in the majority of groups one person is informally subscribed as the leader, who has the responsibility and makes critical decisions for the whole group. In most cases, the choice of this leader is not done on the basis of knowledge concerning avalanches, but rather of secondary characteristics such as sympathy, age or skiing talent (McCammon 2004).

Social facilitation describes the phenomenon that people tend to accept the behaviour of others and see it as good and right (McCammon 2002). The belief that a slope showing traces of others is safe, is a great example which belongs to this heuristic. Also our judgement of risk is different when other people are present. We tend to underestimate risk and expose ourselves to greater risk if we are not alone (McCammon 2004).

Scarcity includes the principal of 'psychological reactance' which describes the occasionally aggressive human behaviour when someone sees himself constrained in his freedom. We do not appreciate prohibitive rules and the 'powder fever' does its bit that every one wants to be the first to leave his traces in fresh new snow. The tendency to overestimate values of prohibited things is omnipresent (McCammon 2002).

McCammon (2004) allocated an exposure score consisting of seven easy recognisable indicators (*avalanche path, recent loading, terrain traps, posted hazard, recent avalanches, thaw instability and instability signs*) to each accident to allow a comparison between accidents with and without signs of heuristics. Single heuristics are considered as well as combinations of them.

As a result, McCammon (2004) proves that in some cases heuristics are present and that the group size as well as the training level have an influence on decisions (see figure 2.3). He concludes that "the six heuristics cues have the power to lure almost anyone into thinking an avalanche slope is safe" (McCammon 2004, 7).

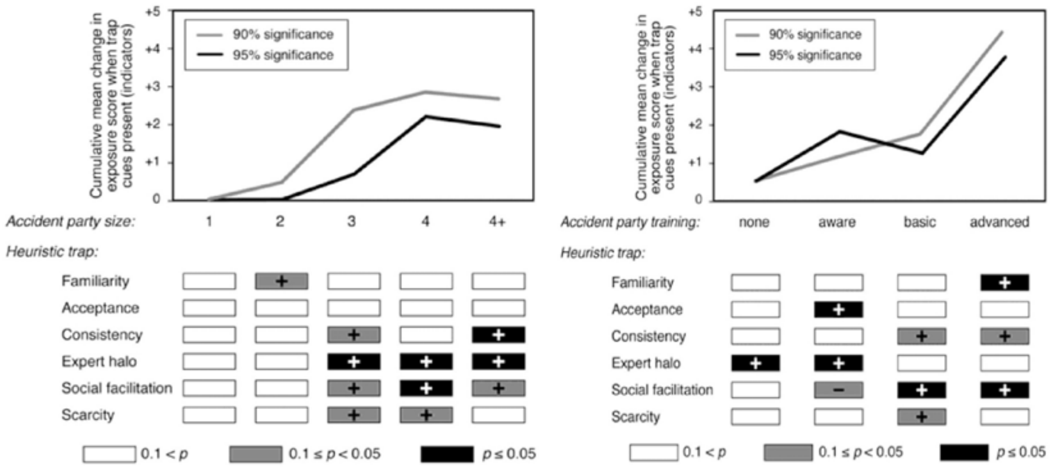


Figure 2.3: Left: Cumulative mean changes in exposure scores for various group sizes | Right: Training levels when heuristic trap cues were present (McCammon 2004, 7).

2.1.2.5 Trigger

The first three factors described in the previous sections can make a situation avalanche prone, because they influence the stress and the strength of the snow pack. A slope is stable as long as the stress, which tends to cause rupture in the snow pack, and the strength, which resists the stress, are balanced. As soon as the stress is getting too big or the strength too small, an avalanche can release (Ward 1984b). Ward (1984b) defines avalanches occurring due to an increase in stress as ***direct-action avalanches***, avalanches which release because of loss of strength as ***climax avalanches*** and a combination of both as ***hybrid avalanches***.

In order for an avalanche to actually release, a trigger is often needed to either increase the stress or reduce the strength in the snow pack. Again, different factors can act as triggers. In most cases, ***internal triggers*** lead to a loss of strength, for example when a weak layer collapses or when meltwater reduces the cohesion of layers (Barton & Wright 2000). ***External triggers*** mostly increase the load on the snow pack and therefore lead to additional stress. This can be caused by new snow, rain or human impact. Therefore "there is no doubt that large parties apply a very considerable additional load to the anchors of a slab" (Barton & Wright 2000, 48). In fact, a person adds an additional load to the snow cover at an area of about two square meters. The more people are on the snow surface, the larger the load and therefore the stress on the snow cover (Harvey *et al.* 2012).

An other classification of avalanche triggers distinguishes between ***natural*** (cornice fall, earthquake, ice fall, rock fall) and ***artificial*** (explosives, vehicle, human, wildlife) ones (Greene *et al.* 2010a).

2.2 Data Related to Avalanches

When working with avalanche data, not only information about the avalanches themselves is important, but also information about the circumstances of the avalanche release, including the factors described in the previous sections.

2.2.1 Acquiring Avalanche Data

For a long time, avalanches have been described and a great amount of avalanche data is available worldwide. One part of these data describes the avalanches themselves. Attributes such as *date, time, observer, avalanche type, trigger, size, dimensions (slab thickness, width, vertical fall)* and *the location of starting zone or terminus* are recorded. The second part of data describes the circumstances of the avalanche release, including the path characteristics and terrain properties *(observation location, aspect, gradient, elevation)* on the one side and the snow properties on the other side (Greene *et al.* 2010a). Path characteristics can either be estimated and measured in the field using a compass, map and further hand-hold instruments or they can be extracted from a map or DEM.

2.2.2 Collecting Snow Cover Data

Information about characteristics of the snow cover is often obtained by digging a snow profile which is "a record of the layer sequence and of the individual layers' properties" (McClung & Schaerer 1993, 141). Snow profiles can be used for several reasons, but their information is restricted to one location and an extrapolation has to be done with great care. According to McClung & Schaerer (1993) *snow temperature, snow hardness, layer boundaries, snow grain form and size, free water content* and *snow density* normally are measured for several intervals or for every distinct layer in the profile. Additionally, stability tests can be conducted to assess how stable a snow cover is.

In the following sections, those properties relevant to the used data in this thesis are presented in detail.

2.2.2.1 General Properties

Snow temperature is normally measured at intervals of 10 cm and temperature gradients are calculated. An accuracy of 0.5 °C can be reached with most thermometers (Greene *et al.* 2010c). Generally, the surface layer of the snowpack has lower temperatures than the deepest layer which is influenced by ground temperature. The closer a layer is to the surface, the higher the influence of the weather variability above the surface. In maritime climates, deep snow packs are usually of relatively warm temperatures and gradients are small. In continental climates, shallow snowpacks show much higher temperature gradients and snow temperatures are lower (McClung & Schaerer 1993).

As long as the snow is dry, snow temperature is not a strong sign for instability (Wiesinger & Schweizer 2001), but snow temperature and gradients have an effect on the velocity of processes that take place in the snow cover, described by Harvey *et al.* (2012) and McClung & Schaerer (1993):

- **< -5 °C**: weaknesses develop slowly and persist for long times. Cold snow tends to be stiffer and more brittle than warmer snow.
- **-5 °C to -1 °C**: rounding and sintering processes are fast.
- **-1 °C to 0 °C**: rapid strengthening of the snow cover but sudden loss of strength when the temperature gets near the melting point due to the influence of melting water. An isothermal snowpack is stable as long as temperatures stay below 0 °C.

- **>10 °C/m**: snowpack is unstable, strength decreases, facet formations take place.
- **<10 °C/m**: strength increases, depending on the characteristics of the snow crystals such as size and form.

Snow hardness describes the “resistance to penetration of an object into snow” (Fierz *et al.* 2009, 6) and can be measured either with rammsondes or by applying the hand hardness test where five classes of snow hardness are distinguished (Barton & Wright 2000, Greene *et al.* 2010c, SLF 2013):

- **F**: a gloved fist can be pushed into the snow layer.
- **4F**: four gloved fingers can be pushed into the snow layer.
- **1F**: one gloved finger can be pushed into the snow layer.
- **P**: a pencil can be pushed into the snow layer.
- **K**: a blade of a knife can be pushed into the snow layer.

Additionally, **I** is sometimes used to describe layers which are harder than a blade of a knife and consist of ice. The harder a layer the higher is the strength of the snow. The first two classes (F and 4F) are generally seen as weak layers (SLF 2013, Wiesinger & Schweizer 2001). The bigger the hardness difference between two adjacent layers, the higher the tendency that the bonding between these layers is loose and that a shear surface might be present (SLF 2013). When the hardness is assessed for each layer, a hardness pattern of the profile can be established. The SLF (2013) distinguishes between ten different shapes of hardness profiles showing various stabilities.

Snow crystal forms can be classified according to the *International Classification for Seasonal Snow on the Ground* where a distinction between the following main classes is made (Fierz *et al.* 2009, Greene *et al.* 2010c):

- Precipitation particles
- Decomposing and fragmented particles
- Rounded grains
- Faceted crystals
- Depth hoar
- Surface hoar
- Melt forms
- Ice formations

Each of these main classes has a variety of subforms (see figure 4.7 for details). Some crystals are intrinsically unstable, especially when they are angled or elongated. Table 2.4 summarises crystal forms that are mentioned by several authors to be unstable. As stabilising crystal forms are melt-freeze crystals and sometimes ice lenses seen and also rime and graupel are only rarely observed in weak layers, with graupel sometimes acting as weak layer shortly after deposition (Wiesinger & Schweizer 2001). McClung & Schaerer (1993) mentions additionally rounded grains as stable crystal forms.

Barraclough *et al.* (2010) analysed typical Scottish snow crystal forms in field tests and found the following six major snow types: powder snow, windslabs, graupel, wet snow, meltcrusts and icecrusts.

Snow crystal size is recorded in [mm], describing the average maximal extension (Greene *et al.* 2010c). Generally, the following classes are distinguished: *very fine* (< 0.2 mm), *fine* (0.2 mm - 0.5 mm), *medium* (0.5 mm - 1.0 mm), *coarse* (1.0 mm - 2.0 mm), *very coarse* (2.0 mm - 5.0 mm) and *extreme* (> 5.0 mm) (Fierz *et al.* 2009). The larger the grains, the bigger the air spaces in between, the lower the number of bondings and the weaker the snow (Wiesinger & Schweizer 2001).

Liquid water content can be measured by forming a snowball, squeezing it and recording the reaction of the snow. The following classification is used (Greene *et al.* 2010c):

- **No snowball**: it is difficult to form a snowball. Snow grains adhere little to each other. Snow temperature is mostly below 0 °C.
- **Snowball**: snowballs can easily be formed. Water is not visible in the snow, also with a 10x magnification. Snow grains stick together when pressed. Snow temperature is at 0 °C.

Table 2.4: Crystal forms mentioned to be unstable by different authors.

Crystal Form	Author
Precipitation Particles	Barton & Wright (2000): **Graupel** Diggins (2011): **Needles, Stellars or Dendrites, Graupel** McCammon (2002) McClung (1993): **Columns, Needles, Plates, Graupel** Ward (1980): **Hail** Ward (1984b): **Hail**
Decomposing, Fragmented Particles	McCammon (2002)
Rounded Grains	-
Faceted Crystals	Diggins (2011) Föhn (1992) McCammon (2002) McClung (1993) Wiesinger (2000)
Depth Hoar	Barton & Wright (2000): **Cup Crystals** Föhn (1992) McCammon (2002) McClung (1993) Wiesinger (2000)
Wet Grains	McClung (1993) Ward (1980): **Slush**
Feathery Crystals (Surface Hoar)	Diggins (2011) Föhn (1992): A**ged Surface Hoar** McCammon (2002) McClung (1993) Ward (1984b) Wiesinger (2000)
Ice Masses	Barton & Wright (2000) McCammon (2002): **Ice Lenses** McClung (1993) Ward (1984b): **Ice Layers**
Surface Deposits and Crusts	Barton & Wright (2000): **Rime** McCammon (2002)

- **Wet glove**: when squeezing a snowball, water cannot be pressed out but the glove gets wet. Water is visible with 10x magnification. Snow temperature is at 0 °C.
- **Water drops**: when squeezing a snowball, water drops out. Some air is still present in confined pores. Snow temperature is at 0 °C.
- **Slush**: the snow is flooded with water. No or only little air is present. Snow temperature is at 0 °C.

The liquid water content has not much influence on snow stability as long as the snowpack is not isothermal. Wet snow tends to be weak as the snow crystals are surrounded with water reducing the strength of bondings (Wiesinger & Schweizer 2001).

Snow density is recorded by dividing the volume of a certain snow amount by its weight (Greene *et al.* 2010c). The density is not a direct sign for instability and in many cases, hardness or grain size difference are more clear indications. However, in general can be stated that layers tend to weaken with lower densities (Wiesinger & Schweizer 2001).

Layer thickness of the weak layer, of the depth of the weak layer below the surface and of the whole snow cover influences the snow stability. Generally “a snowpack with many thin layers is rather more unstable than a snowpack that only consists of a few, relatively thick layers“ and the closer a weak layer to the surface is, the more prone it is to triggering (Wiesinger & Schweizer 2001, 5).

2.2.2.2 Stability Issues

A further possibility to identify weak layers are stability tests, for example the Rutschblock Test, Shovel Shear Test (Barton & Wright 2000, Föhn 1992, Schweizer & Jamieson 2010) or Collapse Test, Tilt Board Test and Shear Frame Test (McClung & Schaerer 1993) to only mention some. Those tests important to the data in this thesis are shortly described.

In a **Rutschblock Test**, a block of snow, representing an area that can be affected by the load of a skier, is separated on a representative slope. A person approaches the block from above and conducts the steps listed below (McClung & Schaerer 1993). It can be differentiated between *Walking Rutschblock Test* (when a person conducts the test without skies, size of block: 1m by 1m) and *Ski Rutschblock Test* (when a person is on skis or snowboard, size of block: 2m by 1.5m) (Barton & Wright 2000, Greene *et al.* 2010c):

- **1**: failure of the block when cutting the back side.
- **2**: failure when a person approaches the block from above.
- **3**: failure when a person is standing on the block.
- **4**: failure when a person standing on the block makes a rapid knee bend.
- **5**: failure when a person standing on the block makes a small jump.
- **6**: failure when a person standing on the block makes a bigger jump.
- **7**: no failure.

Numbers 1 to 3 indicate poor stability, numbers 4 and 5 moderate stability and numbers 6 and 7 good stability (Barton & Wright 2000, Greene *et al.* 2010c, McClung & Schaerer 1993, SLF 2013). The way in which the block fails can give additional information, whereas a distinction between the *'whole block'*, *'most of the block'* and only an *'edge of the block'* is made (Greene *et al.* 2010c).

With the **Shovel Shear Test**, a 30 cm x 30 cm column of snow is cut with the shovel to a depth below the expected weak layer. Pressure is applied to the shovel in the backcut of the column and the following resulting classes are distinguished (Greene *et al.* 2010c, McClung & Schaerer 1993):

- **Collapse**: block collapses when cut.
- **Very easy**: column falls during cutting or when inserting the shovel.
- **Easy**: column fails with a very low shovel pressure.
- **Moderate**: column falls with a moderate shovel pressure.
- **Hard**: column falls with a firm, sustained shovel pressure.
- **No shear**: no shear failure.

To assess the general stability of a snow cover, several measurements have to be combined as they are often subjective and only provide information for one specific location. "The results of any stability test must be coupled with snowpack and weather histories, shear quality, snow structure, and other observations before the stability can be assessed" (Greene *et al.* 2010c, 44). The Rutschblock score is a strong indicator, but Schweizer & Jamieson (2003) show that other profile parameters have a similar power as well. Different approaches to integrate the variety of criteria into one stability measure or to evaluate which attributes are most important and have the greatest power exist.

A **Shear Quality Index** was developed to use in combination with the above presented stability tests and can give additional information about the general stability of a snow cover. The following classes are distinguished (Greene *et al.* 2010c):

- **Q1**: clean, planar, smooth and fast shear surface.
- **Q2**: average shear, mostly smooth but not as fast as Q1, some irregularities but not as much as Q3.
- **Q3**: non-planar, uneven, irregular and rough shear surface.

Johnson & Birkeland (2002) assessed how shear quality can be integrated into stability test results and found that this additional information can be very important, especially for cases in which a generally good stability is found but a Q1 shear quality is present.

According to the SLF the following measures should be used to identify the stability of a snow cover: *general hardness pattern, grain shape, size, difference in grain sizes, hardness, difference* *in hardness, snow depth and thickness of sliding layer, temperature and liquid water content* (SLF 2013). They developed a measure called **Nietentest**, that combines these different variables and includes the *Rutschblock score, type of fracture in Rutschblock test, grain size difference, grain size, hardness difference, hardness, grain shape and depth* (Schweizer 2007). If maximally two of these variables show instability, the snow cover is seen as having more or less good stability and no pronounced weak layers. With three or four instability signs, the possibility for a weak layer exists and with five or six instability signs, the snow cover is seen as rather unstable with a high probability for a critical weak layer (Schweizer 2008). Thereout, the following classification for snowpack stability is proposed by Schweizer (2007):

- **Low stability**: clearly defined weak layers. Rutschblock: 1, 2, 3.
- **Medium stability**: weak layers exist. Rutschblock: 4, 5, (3 possible).
- **Good stability**: no recognisable weak layers. Rutschblock 5, 6, 7.

Other authors tried to assess the general stability of a snow cover as well. Schweizer & Jamieson (2003) used the combination of the following variables as stability indicator: *Rutschblock score, grain size, layer hardness, difference in grain size* and *difference in hardness*. They summarise that "soft slabs, large and persistent grains, large difference in grain size and hardness and low Rutschblock scores" are indicators of snow instability (Schweizer & Jamieson 2003, 243). McCammon & Schweizer (2002) combine *fracture depth, weak layer thickness, hardness difference, grain type* and *grain size difference* as signs for instability and state that none of these parameters can be used as a reliable single measure but that the threshold-sum is a good indicator for snow cover instability. Wiesinger & Schweizer (2001) propose the following attributes indicating instability: *Rutschblock score, hardness, presence and type of weak layers, grain type* and *size*.

Some of the authors mentioned above found threshold values for their criteria as well. Table 2.5 shows the thresholds used by the SLF and McCammon & Schweizer (2002) mention the following:

- **Grain size difference**: >= 1 mm
- **Hand hardness difference**: >= 1
- **Depth of weak layer**: < 1 m below the snow surface
- **Layer thickness**: < 10 cm
- **Layer**: consisting of persistent crystals

Table 2.5: Critical thresholds of different snow pack properties for the instability of the snow cover (SLF 2013).

Variable	"Critical value range"
Rutschblock score	< 4
RB: type of fracture	Entire block
Grain size difference	≥ ¾ mm
Grain size	≥ 1¼ mm
Hardness difference	≥ 2 steps on the hand hardness index
Hardness	≤ 1-2
Grain shape	faceted, depth hoar or surface hoar
Depth	< 1 m

2.2.3 Recording Weather Data

To enable an even more rigourous data analysis, weather attributes before as well as at the time of an avalanche release can be obtained. Often, this data is acquired manually and directly in the field in combination with snow profiles and can consist of the following attributes: *observation location, elevation, date, time, observer, sky conditions, precipitation type, rate and intensity, air temperature and trend, relative humidity, pressure and trend, snow temperature, surface penetrability (ram, foot or ski penetration), form and size of surface snow, total snow depth, 24-hour new snow depth, 24-hour new snow water equivalent, snow density, rain, 24-hour liquid precipitation, wind direction, wind speed, maximum wind gust* and *blowing snow* (Greene *et al.* 2010b). Many visual formations giving clues about wind speed and wind directions such as ripples, sastrugi, cornices, drifts, deep loose snow, rime accumulations or trees and poles can be used additionally (McClung & Schaerer 1993). Penetration indices can give an indication of how much loose snow is available for the formation of avalanches and for wind drift (McClung & Schaerer 1993).

If possible, manual data can be enhanced with continuous data measured at automated weather stations which mostly include information about *wind speed, wind direction* and *air temperature*. According to Greene *et al.* (2010b, 5) data of automated weather stations "allow observers to fill in the periods between manual observations", but they should be used "to augment and not replace manual observations". If the data include careful records of the source and type of measurement, manual and automated data can be combined. "With several years of data, usually good correlations can be achieved between anemometer readings and wind at a specific location" (McClung & Schaerer 1993, 158/159). Jomelli *et al.* (2007) mention that climatic data in general may come from stations away from the exact location of an avalanche event, so that meteorological parameters calculated from any station are not strictly identical to those at the avalanche locations. Further "mountain peaks and ridges are not always the best locations because their positions may be too exposed" (McClung & Schaerer 1993, 158/159).

2.3 Analysis of Avalanche Data - Example Studies

Data as mentioned above are acquired in many studies worldwide and analysed to achieve different aims. Some studies summarise avalanche activity, give an overview of past events, extract characteristics of avalanches, say something about their physical properties, distributions, frequencies and return periods, other studies use the data for establishing hazard zones or as basis for forecasting.

2.3.1 Overview of Avalanche Activity

Overviews of avalanche activity including a descriptive summary of characteristics of avalanches have not been done often.

Stoffel *et al.* (1998) analysed an avalanche data set from a fourteen year period from the region of Zuoz in the Engadine valley in the Swiss Alps. They aimed to assess where avalanches occur, how big they are, how often they release and what the related conditions look like. Between 1982 and 1996, about 1100 avalanches have been mapped. Further, daily snow and weather information was available, where the following variables have been measured: *new snow depth, snow depth, weather type and intensity, wind direction and speed, air and snow temperature, snow-surface conditions, ram penetration depth* and *water equivalent of new snow*. Additionally, bimonthly data of a snow profile was available. Stoffel *et al.* (1998) digitised avalanches, defined potential starting zone areas and summarised avalanche frequencies, aspects of starting zones and avalanche sizes. Results show that during the fourteen years of observation "a wide variety of contributory factors" (Stoffel *et al.* 1998, 329) could be found and "that half of the potential starting zone avalanched at least once" (Stoffel *et al.* 1998, 335).

In the Catalan Pyrenees, a database describing the avalanche activity over twenty years has been established resulting in a map showing potential avalanche areas. Oller *et al.* (2006) aimed to analyse this data set by importing it into a GIS by considering information about starting and runout zone elevation, aspect and avalanche size. Results are shown in box plots and an aspect diagram.

Summarising, the study has shown that “climate has more influence than aspect in avalanche starting conditions” (Oller *et al.* 2006, 312).

Hägeli & McClung (2003) analysed avalanche data of the Columbia Mountains in Canada and classified each winter between 1980 and 2000 either as continental, maritime or transitional one and described the avalanche activity with focus on weak layers, including an analysis of their distribution and variability. An *Avalanche Activity Index* is used which consists of the sum of the sizes of avalanches per day. Results showed that weak layers play a role in 16% of the natural slabs and that approximately twice as many weak layer slabs occur in continental than in maritime winters.

More studies using avalanche data exist, however, with different aims.

Eckert *et al.* (2007) share the opinion that it is possible to encounter the sparsity of avalanche data by using spatial analysis. On the scale of township, the authors have done interference and predictive sampling steps using a hierarchical Bayesian Model and Monte Carlo Simulation. They show that 60% of the total variability can be explained through the spatial structure of the data. Because avalanche data belongs to the group of rare discrete events, a Poisson Model can summarise the local avalanche activity.

Laternser & Schneebeli (2002) used avalanche data to define a regional *Avalanche Activity Index* in order to look at long term changes and assess the quality of data coming from 84 Swiss avalanche observation stations. For each of the avalanches which occurred during the 50 year period in the Swiss Alps, the *triggering mechanism, avalanche type, slope direction and elevation of starting zone, number, size* and *impact* was available (Laternser & Schneebeli 2002).

Casteller *et al.* (2011) tried to define years with major avalanche activity in the Patagonian Andes using dendrochronology and correlate them to climatic records and atmospheric patters. Years in which the average *Event Index* exceeded the mean of a long term period were seen as years with major avalanche activity. Monthly precipitation and temperature measured at meteorological stations were compared to mean values and it was looked at anomaly maps of geopotential heights and wind vectors. For snowfall and wind attributes no data was available. Results show that the “total monthly precipitation during the three snowiest months was significantly greater than for years without large avalanche activity” and that “atmospheric patterns show features typically observed during the cold phase of El Nino [...] resulting in both higher precipitation and stronger winds” (Casteller *et al.* 2011, 68).

2.3.2 Pattern Analysis

In the Oxford Dictionary the term ’pattern’ is defined as “an arrangement or sequence regularly found in comparable objects or events” and Watanabe (1985) in Jain *et al.* (2000, 3) describes patterns “as opposite of a chaos; it is an entity, vaguely defined, that could be given a name”. In this thesis, patterns are used as description for situations which lead preferably to avalanches. These situations can be described with different factors and their specific characteristics at the time of avalanche release.

The reasoning for the application of patterns in avalanche research is formulated by Nairz (2009, 389) the following: “it is no coincidence that avalanche accidents occur at similar times at similar places / spatial distribution during analogue conditions” as similarities in the structure of the snowpack and the weather “lead to similar avalanche-relevant situations (patterns)”.

Patterns are a good way of communicating complex processes to the public as they summarise a huge amount of data, processes and interpretations. Patterns are easy understandable, can be learned and remembered without huge effort and therefore play a very efficient role. Harvey (2008) summarises the motivation for pattern analysis with the necessity of focusing on important facts to enable right decisions in a short time when being in avalanche terrain. Because of the ability of our brain to recognise characteristics we have already met before, experience is very important and we can learn from the past when remembering similar situations. Also Jain *et al.* (2000) mentions that humans are best pattern recognisers.

Different services have already established avalanche patterns. One example is the Lawineninformationsdient in Tirol, Austria. Nairz and Mair have published ten patterns which account for at least 98% of all hazardous situations during the winter season (lawinenwarndienst tirol 2013, Mair & Nairz 2010, Nairz 2008):

1) The second snowfall
2) Sliding snow
3) Rain
4) Cold following warm / warm following cold
5) Snowfall after a long period of cold
6) Cold, loose new fallen snow and wind
7) Snow-poor zones in snow-rich winters
8) Surface hoar blanketed with snow
9) Graupel blanketed snow
10) Spring time scenario

These patterns consider the snow pack as well as related weather conditions and not only represent one winter season, but situations which occurred in several different winters. The patterns show temporal as well as spatial differences and also the danger potential differs (Nairz & Mair 2010).
The SLF (Schnee und Lawinenforschungsinstitut der Schweiz, Harvey 2008, Harvey *et al.* 2012, Harvey 2013, SLF 2013) presents four main avalanche patterns, in which each can be characterised by slightly different conditions. Nevertheless, this simple structure helps to focus on the most important avalanche building factors. To allow the relation of the patterns to the terrain and specific slopes, they have to be considered always in combination with the two questions “What is the most important danger?” and “Where does this danger probably occur?” (Harvey 2008).

New Snow In new snow situations, snowfalls have occurred during the last one to three days. Important questions which have to be asked concern the critical new snow amount (see section 2.1.2.2), the snow cover on which the new snow has fallen and the time of the most severe settlement processes.

Drifted Snow Wind transports loose snow and builds dangerous areas of drifted snow. Important related factors are gradient and aspect of the slope, the underlying snow cover, if new snow or old snow has been transported and if new snow has fallen on the drifted snow.

Old Snow Old snow situations are present when for at least three days no crucial changes in precipitation, wind or melting processes have occurred. Weak layers can persist in old snow covers. The situation is especially pronounced when an alternation of little snowfalls and cold temperature periods has been present. Factors which have to be considered are critical layers in the snowpack, the variability of the layering and the happenings since the last snowfall.

Wet Snow In wet snow situations, water is influencing the snowpack and the snow gets heavier.

The avalanche patterns presented by the Lawinenwarndienst Tirol and the SLF are two examples of how a result of a pattern analysis can look like. They differ in many ways, one reason being the different locations to which they are aimed. The patterns of the Lawinenwarndienst Tirol should be applied in Austria, the patterns of the SLF in the Swiss Alps. In Scotland, conditions differ and, therefore, these patterns cannot be applied directly and have to be established considering the available Scottish data. Some example studies exist, where authors aimed to find patterns in avalanche data.

Stoffel *et al.* (1998) tried to relate weather and snow cover data to avalanche activity. They considered *three day sum of new snow depth* and compared the values on avalanche days. Using six example days, conclusions on other factors such as temperature, wind speed and direction, ram profile and hazard danger level are drawn. The results show that “even for large amounts of new snow, the new-snow depth is not sufficient to explain the extent of avalanche activity. [...] The temperature evolution during, and the snow-cover conditions prior to, the snowfall seem to be most decisive” (Stoffel *et al.* 1998, 335).

Jomelli *et al.* (2007) assessed relationships between avalanche occurrences and meteorological parameters in the French Alps for the time period between 1978 and 2003. *Precipitation (in mmWE) on the day and on days before, minimum, maximum and mean air temperature* as well as *thermal* *amplitudes of the day and days before* are considered. Further, variables such as *total winter precipitation, number of times with two and three consecutive rainy days, freeze-thaw alterations or mean air temperatures more than 1/2, 1 or 2 standard deviations above or below the winter mean* are tested on an annual basis. Results show that "the probability of avalanche occurrence correlated primarily with precipitation and most often with total precipitation for the 3 days preceding a given avalanche event" and that "mean temperature on the day of avalanche events seemed a key variable. [...] Wind direction and speed may also have a strong influence on avalanche occurrence, but these parameters were not tested" (Jomelli *et al.* 2007, 189).

McCollister *et al.* (2003) developed an interactive database tool called *GeoWAX* using a nearest neighbour approach to combine meteorological and avalanche data. For a data set over 23 winter seasons at the Jackson Hole Mountain Resort, Wyoming in the US, the authors looked exemplarily at the three variables *new snow, wind speed* and *wind direction*. Input parameters for the nearest neighbour model consisted of 34 variables, including *new snowfall, snow water equivalent, total snow depth, minimum and maximum temperatures, wind speed and wind direction from a summit station*, as well as subjective parameters such as *snow amount available for wind transport* or *daily warming* (McCollister *et al.* 2003). Avalanche data consisted of the *date and time of the avalanche release, avalanche type, trigger mechanism, size, depth, sliding surface* and *slide path name*. Results show that "an increase in new snowfall leads to an increase in the avalanche probability at all scales" and that wind speed and wind direction show different influences when looking at a single avalanche path or the entire area which is a sign for the high local variability (McCollister *et al.* 2003, 306).

The study of Birkeland *et al.* (2001) aimed to set extreme avalanche days in relation to atmospheric circulation patterns. Four sites in the West of the United States were examined on a daily time-scale. An *Avalanche Hazard Index* was defined by the sum of the squared avalanche sizes for each day. The upper 10% of the days were then used as extreme avalanche days. For the avalanches, the number and size for each day were available and the climatic data included *maximum and minimum temperatures, total snow depth, new snowfall, snow water equivalent* and *rainfall* (Birkeland *et al.* 2001). Additionally, *two and three day new snowfall* and *snow water equivalent* were calculated to account for storms. A varimax rotated PCA was run on time-attribute data matrices which included the original variables as well as the additionally calculated ones. The component scores for each day were analysed and days with scores greater than plus or minus 1 were seen as abnormal conditions. The component scores of the extreme avalanche days were analysed to extract the most important climatic variables. As a last step, composite-anomaly maps at the 500 hPa level were created to relate atmospheric patterns at a larger scale and see where differences to average conditions occur. Results show that 93% to 96% of the variance in the climate data can be explained by the components found through the PCA (see figure 2.6). The component including one day snow and snow water equivalent as well as the two and three day snow and snow water equivalent have been found to be most important. This shows that short term snow fall and snow water equivalents have a high influence, in some cases also prolonged storms. Temperature in all four study areas was judged as being less important.

Table 2.6: The total variance of the original climate-data matrix explained by the PCA, and the amount of variance explained by each component (Birkeland et al. *2001, 137).*

	Bridger Bowl	*Jackson Hole*	*Alta*	*Taos*
Principal components	2 and 3 d SWE (23%) 1 d snow and SWE (20%) Max./min. temp. (20%) 2 d snow and 3 d SWE (19%) Snow depth (11%)	2 and 3 d snow and SWE (38%) 1 d snow and SWE (23%) Max./min. temp. (21%) Snow depth (11%)	2 and 3 d snow and SWE (36%) 1 d snow and SWE (26%) Max./min. temp. (21%) Snow depth (11%)	2 and 3 d snow and SWE (36%) 1 d snow and SWE (26%) Min. temp. (14%) Snow depth (11%) Max. temp. (9%)
Total variance explained by PCA	93%	93%	94%	96%

Similar approaches were used by Esteban *et al.* (2005b) who applied a PCA to define the number of clusters and the initial cluster centres of the following k-means CA. The methods were applied to data of Andorra, Pyrenees, and aimed to cluster days with heavy snowfall (more than 30 cm in 24 hours) and relate them to synoptic circulation patterns. The PCA resulted in six principal components that explain 89.7% of the total variance and six severe atmospheric circulation patterns resulted (Esteban *et al.* 2005a).

García *et al.* (2008) applied a PCA to data of the Eastern Pyrenees and related avalanche activity to six atmospheric patterns which could explain 94% of the total variance. Mock & Kay (1992) found even three components which explain 85% of the total variance by applying a PCA to avalanche, climate and snowpack data of the Western United States.

Kleemayr *et al.* (2000) applied a variety of methods on avalanche data to establish a prognosis tool. One of these methods included a CA with which the data was summarised into groups. The following variables were used as input: *air temperature, air temperature difference to the day before, air temperature difference to two days before, sum of new snow over the last two days, snow depth, sink depth, snow temperature* and *snow temperature two days before*. Eight clusters resulted with the following characteristics: 1) cold, little new snow, decrease in temperature; 2) cold, little new snow, middle strong decrease in temperature; 3) no new snow, strong warming with deep temperatures; 4) much new snow, deep temperatures, decrease in temperature; 5) no new snow, small warming, air temperature around 0 °C; 6 and 7) cold, little new snow, strong temperature decrease; 8) no new snow, strong warming at 0 °C.

Harvey *et al.* (2002) and Signorell (2001) analysed avalanche accident data over a time period of 30 years between 1970 and 1999 in the Swiss Alps to combine them with different snow and weather parameters such as *amount of new snow, total snow depth, snow temperature, wind speed* and *air temperature* (Harvey *et al.* 2002). Avalanche days were defined as days on which at least four avalanche accidents occurred. Snow and weather data were transformed in the following nine variables (Harvey *et al.* 2002): *sum of new snow and wind speeds over the last seven, three and first four of the last seven days, percent of total snow depth in relation to long term average, difference in air* and *snow temperature between the avalanche day and the day before*. The mean of these nine variables has been assigned to each avalanche day and a k-means CA was conducted resulting in five different clusters of avalanche days (Harvey *et al.* 2002):

Cluster 1: rise in temperature and little new snow
Cluster 2: high amount of new snow and low temperature
Cluster 3: unstable snow cover, mean amount of new snow/winds, low temperature
Cluster 4: strong winds
Cluster 5: catastrophe: much new snow, mean winds, unstable snow cover

Weather and snow situations before avalanche days were not exceptional in the majority of cases. At most avalanche days, a danger level of 'Considerable' was published. In the majority of cases, the combination of several factors led to an accumulation of avalanche accidents. Often, the weather was nice and frequently, avalanche days fell on Sundays or holidays. Under these conditions, many people are out in the mountains and the potential for a high damage increases (Harvey *et al.* 2002). Summarising, considerable amount of new snow has to be present on an avalanche day. In 20% of all avalanche days, neither a considerable amount of new snow nor strong winds occurred, but the temperature has risen (Harvey *et al.* 2002).

3

SAIS and Study Area

Avalanches in Scotland are special, for several reasons. Moss (2009, 628) describes the avalanche situation the following: "avalanche hazard in Scotland is characterised by rapidly changing weather, rugged topography and a snowpack shaped above all by wind." Barton & Wright (2000, 133) state that "winters in Scotland vary enormously from one season to another. [...]. Almost the only generalisation it is possible to make is that it is unlikely that consecutive winters will be similar" and the Met Office describes on their webpage: "the UK is well known for the variability of its weather - from day to day, season to season, year to year and place to place" (MetOffice 2013c). Also as advertising these special weather conditions are used, for example on the webpage of Cairn Gorm: "the ice and snow conditions can change daily. It can change from powdery soft snow to slush to neve in a matter of days. That's what makes coming here such a challenge; you are entirely at the mercy of the weather, whatever it decides to do" (cairngorm mountain 2013).

The Scottish climate is influenced by the Gulf Stream as well as by the continental climate of Europe. Generally, higher temperatures, less precipitation, weaker winds and more sun characterise the weather in the east and south of Scotland in comparison to westerly and northerly regions. Due to the small scale variability of the topography and land use, weather conditions are highly variable (MetOffice 2013c).

Considering avalanche activity, this leads to important differences compared to other mountain ranges. Looking for example at the Alps, Ward (1980) summarises that depth hoar develops only rarely, snow density and penetration values are higher and that ice layers occur much more frequently. Further, snow temperatures are generally higher and windslabs play a crucial role (Ward *et al.* 1985). Also snow crystals show special characteristics with a "prevalence of rounded and decomposed grains, often interspersed with crusts or ice lenses and hardened by wind or freeze-thaw cycles" (Barraclough *et al.* 2010, 367).

This chapter presents relevant aspects considering Scottish avalanches. The Scottish history of winter sports is introduced, the SAIS is presented, background information about the study areas is provided, the Scottish climate is broadly characterised and most important steps in research related to Scottish avalanches are summarised.

3.1 Winter Sports in Scotland

First ascents at Ben Nevis date in the late 19th century, in the Northern Cairngorms in the beginning of the 20th century. One of the first popular climbs was achieved by the three guys Goodeve, Russel and Robertson who climbed the Central Gully in Coire an-t-Sneachda in 1904 (highland guides 2013). In early years, the climbing took place preferably in gullies and Scottish winter cliffs were mainly used “as a training ground for Alpine mountaineering” (Patey 1960, 186). Around 1930, winter mountaineering experienced a revival when “all snow conditions were regarded as climbing conditions. It merely became a question of adapting the technique to meet the prevailing conditions” (Patey 1960, 186). In the forties, the first two climbing schools opened, using Glencoe and Lochnagar as primary climbing grounds, and so “Glencoe was the birthplace of modern Scottish ice climbing” (Patey 1960, 189).

This is, how Patey (1960) describes the early beginning of the Scottish climbing history. Until today the popularity has even risen and according to Mather (1997) in Joyce (2001, 18) are today “climbers accounting for 30% of all visitors to the Scottish mountains during winter”. But not only climbing has gained on popularity, also a broader variety of winter sports is carried out frequently.

In the middle of the 20th century, the first ski resort opened in Glencoe, followed later by the Nevis Range, Cairn Gorm, Glenshee and The Lecht. Today, skiing is popular not only on pistes, but also in free terrain. Bryden *et al.* (2010, 41) state that “skiing remains important to Scotland’s winter tourism” and the Sports Minister Shona Robinson summarised on 04.01.2011: “this year’s winter sports season has got off to a terrific early start and is already shaping up to be a record-breaker with Cairngorm, Glenshee, the Nevis Range, The Lecht and Glencoe all opening early this year and recording a phenomenal 77,000 skier days since the end of November” (ScottishGovernment 2011).

Walking is today “clearly the most popular nature based activity” (Bryden *et al.* 2010, 42) and the number of people has increased mainly since the mid 20th century (Watson 1984). Also HIE (1996) in Hanley *et al.* (2001, 36) summarise that about 77% of the mountaineers are hillwalkers, 11% rock-climbers, 5.5% ski-mountaineers and 6.5% ski-tourers.

That an increase in the number of mountaineers in general is present show for example visits to Scottish climbing sites or the “increase in the number of people annually registered as completing all 279 mountains over 3,000 feet” (Hanley *et al.* 2002, 3).

3.2 Scottish Avalanche Information Service - SAIS

In 1988, the SAIS started operating in the two areas Glencoe and the Northern Cairngorms. The Scottish Sports Council was responsible for the project and the Meterological Office provided support. In both areas, three observers were engaged to ensure a daily coverage and they had direct contact to the Glasgow Weather Centre. The published snow and avalanche reports were displayed on bulletin boards, in the Meteorological Office’s Mountaincall, in a service called Climline and in the two newspapers *The Harald* and *The Scotsman* (Barton & Wright 2000). At that time, the support from outside was not overwhelming and a wide coverage and prominence was hard to gain. But “gradually, however, as it became apparent that the provision of reports was not intended to threaten the freedom of choice of those going to the hills, and that individuals within the service saw themselves as among the custodians of the Scottish mountaineering tradition, this problem receded” (Barton & Wright 2000, 123).

For forecasting, daily surface snow pits were taken and variables such as crystal size and type, wetness and hardness were measured. Also shovel tests and weekly ram penetrometer profiles were conducted. Already then, strong commitment has been given into the gathering of information but “it has to be said that the recording of this information in those early days, was less than systematic” (Barton & Wright 2000, 123).

In the season 1989/90, Lochaber was added to the operation area, as a reaction to the opening of the Aonach Mor ski area (Barton & Wright 2000, 124). In the same season, the reporting service was computerised and in winter 1993, the SAIS webpage was the first in the world which was able to publish forecasts on a daily basis. The webpage *(www.sais.gov.uk)* is still used today, not only for publishing forecasts and additional information, but it enables also communication with the public (Barton & Wright 2000).

After some problems in managing forecasts for Lochnagar and little satisfying experiments of a limited operation of one observer on weekends, the two areas Southern Cairngorms and Creag Meagaidh were added to the operation area of the SAIS in 1996 (Barton & Wright 2000). The Scottish Sports Council was renamed and became **sport**scotland Avalanche Information Service, in short **SAIS** (SAIS 2013b).

All the time, work on the awareness of the Scottish avalanche problematic and cooperations with academic institutions have taken place and international linkages have been established. In 1990, a model (NXD model of the Swiss Federal Snow and Avalanche Research Institute) was tested as support for the forecasting the first time (Barton & Wright 2000). Later, the nearest neighbour model *Cornice*, which allows to compare conditions with similar situations from past winters, has been used “as an additional information gathering tool” (Purves *et al.* 2003, 354). A variety of further projects are on-going. Contemporary collaborations with the University of Edinburgh in a research project about snow and ice mechanics as well as information exchange with the Met Office and SEPA concerning floodings are taking place (Diggins 2010, Diggins 2011b, Diggins 2012).

For the future, Barton & Wright (2000) state that the international linkages will get even more important and also the operational area could develop. But adding a new area is a big step and the number of accidents have to be comparable to Creag Meagaidh. However, visitor numbers on the webpage show that “more people are now aware of the SAIS service than in previous years, and that the breadth of visits worldwide indicate that interest may go beyond the reach of the service nationally” (Diggins 2010, 4).

3.3 Study Area

The study area is located in the Scottish Highlands between Glasgow, Dundee and Inverness in the Grampian Mountains. The SAIS is currently operating in the five areas ’Northern Cairngorms’, ’Lochaber’, ’Glencoe’, ’Southern Cairngorms’ and ’Creag Meagaidh’[1]. They lie at about 57°N and 4°W, including with Ben Nevis (1344 masl) the highest Scottish mountain. All five areas are well known for different winter sports, a variety of activities are offered and they can be reached by car or public transport. Each area is known for popular summits, which are shown in appendix A.4.

3.3.1 Cairngorms

The Cairngorms are the most easterly located SAIS areas and are divided into Northern and Southern Cairngorms, with the village of Braemar lying in between. The Cairngorms National Park in the Cairngorms Massif is the largest British nature reserve and “the Cairngorms are among the coldest and wildest places in Britain with some of the best rock and winter climbing in its heart” (cairngorm mountain 2013).

The Cairngorms are a plateau with large areas above 1000 masl, many cliffs, buttresses and a number of gullies. Ben Macdui with its 1309 masl is the highest summit (Ward 1980). The Northern Corries, including *Coire an t-Sneachda* and *Coire an Lochain*, belong to the most well-known climbing areas (cairngorm mountain 2013).

On the popular Cairn Gorm (1245 masl) different activities are carried out. A commercial ski resort near Aviemore offers 30 km of pistes and a funicular brings people up to Ptarmigan Top Station at 1097 masl (ski mountain 2013). The Glenmore Lodge, a National Outdoor Training Centre offering courses in diverse outdoor adventure sports, is located at the foot of Cairn Gorm (glenmore lodge 2013).

The Southern Cairngorms include two main locations which are well-known for winter sports. The ski resort Glen Shee is the largest in Scotland today and is spread over the slopes of the four mountains Glas Maol (1068 masl), Cairnwell (933 masl), Meall Odhar (920 masl) and Carn Aosda (915 masl) (ski mountain 2013: Glen Shee). On Lochnagar (1155 masl) is the second popular area with its summit being “one of the most famous Munros, a celebrated, pointed summit high above one of Scotland’s most beautiful corries” (walkhighlands 2013b).

3.3.2 Lochaber

The area Lochaber, lying in western Scotland near Fort William, includes the three very popular summits Ben Nevis, Carn Mor Dearg and Aonach Mor, all belonging to the Nevis Range. Ben Nevis

[1] The following abbreviations for the area names are sometimes used: **NC** for the Northern Cairngorms, **LO** for Lochaber, **GL** for Glencoe, **SC** for the Southern Cairngorms, **ME** for Creag Meagaidh.

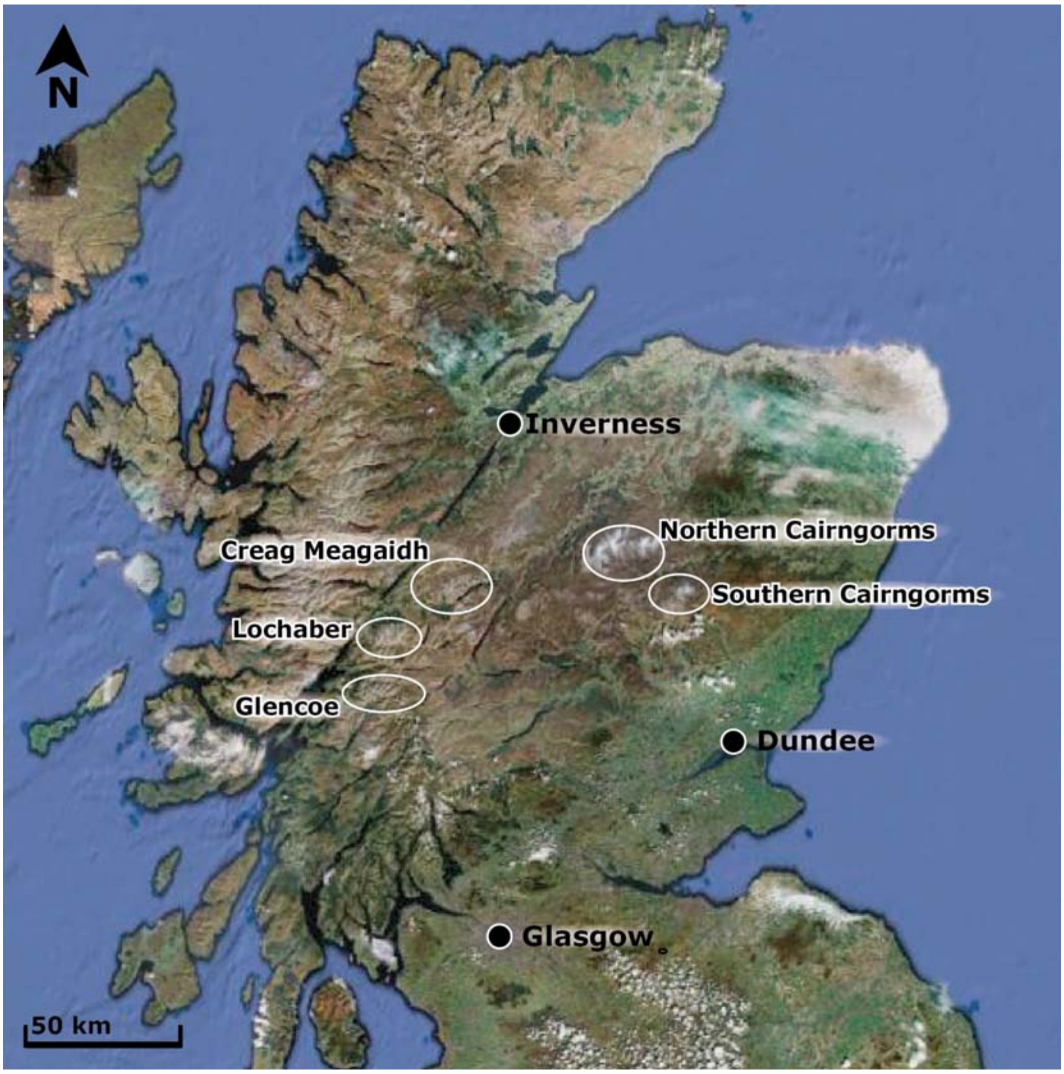


Figure 3.1: The five areas for which the SAIS currently publishes daily forecasts (background: Google Earth).

is with its 1344 masl the highest summit in Scotland and is especially known for ice climbing routes (Moss 2009). For Ben Nevis, a two to three hour walk from the road is needed, but winter climbing is then possible on “a wide variety of aspects, altitudes and difficulties” (Moss 2009, 629). Because of the considerable altitude, snow can hold until April and conditions are very reliable (nevis range 2013).

The newest ski resort in Scotland was opened in 1989 in the Lochaber area. It is the only one using gondolas which bring people to a height of 655 masl at the north face of Aonach Mor which reaches 1221 masl (ski mountain 2013: Nevis Range).

3.3.3 Glencoe

The Glencoe area is the most southerly located one. The main area for winter sports lies south easterly above the village Glencoe, including mountains such as Bidean nam Bian, Buachaille Etive Mor, Creise or Meall a’Bhuiridh. Several ridges and valleys are oriented from south west to north east on both sides of the A82.

The first ski area in Scotland was built in 1956 in Glencoe (ski scotland 2013) and pistes are spread over the slopes of Meall a’Bhuiridh (1108 masl) (ski mountain 2013: Glencoe)

3.3.4 Creag Meagaidh

The summit Creag Meagaidh, a plateau with several ridges reaching 1130 masl, is the heart of the Creag Meagaidh winter sport area on the northern side of Loch Laggan. Most popular is the Coire Ardair between the two easterly ridges, surrounding the Lochan a'Choire with its cliffs and making up the east face of Creag Meagaidh (UKClimbing 2013).

For some people, Creag Meagaidh is ranked directly after Ben Nevis, providing "the best venue in terms of quality and quantity of winter routes" (UKClimbing 2013) and is regarded as one of Scotland's best Munros (Munros 2013).

3.4 Climate

The SAIS areas are, although lying at a northerly latitude, characterised by a temperate, maritime climate. The Scottish climate is influenced by the North Atlantic Ocean and its Gulf Stream, leading to warmer conditions than would be normal at that latitude (Fyffe 2000). Temperatures can rise above the melting point at any time of the year (Barraclough *et al.* 2010, Purves *et al.* 2003) and can increase as high as 15 °C also in winter months (MetOffice 2013c). Not only high air temperatures are possible during the winter, the short distance of not more than 80 km to the sea can also lead to very fast temperature changes "by 10 or 15 degrees within a few hours" as well (Ward 1980, 34).

Generally, temperatures vary on a seasonal as well as on a daily time scale and are lowest at the time of sunrise and highest about two to three hours after noon. Minimum temperatures are generally less than -4 °C, are frequently lower than -8 °C (Ward *et al.* 1985) and can fall below -10 °C in January and February (Ward 1980), which are the coldest months during the Scottish winter (MetOffice 2013c). These very deep temperatures arise when cold Arctic air is transported southwards and influences the Scottish climate as well (Fyffe 2000).

Often, these cold air masses collide with warm southerly ones and are a major cause of storms and rainfalls (Dawson 2009). Storms are frequent in Scotland, often of considerable force and can hold over several days. According to Buchan (1890, XXXVII) "the longest continued storm extended over seven days without intermission" and showed average wind speeds of 75 mph, average maximum wind speeds of 88 mph and minimum average wind speeds of 65 mph. Also Fyffe (2000, 8) mentions the long persistence of strong winds: "bad weather can occur at any time, with winds of over 100 mph being common as are galeforce winds which may blow continually for days at a time". As mentioned, very high wind speeds are characteristic features of Scottish storms as well, which has also Buchan (1890, xlix) stated: "considerable and specially rapid changes from the normal differences of temperature and pressure are frequent concomitants and precursors of storms, but more particularly of destructive winds during storms". Maximum wind speeds occur between December and February (MetOffice 2013c) and therefore during the coldest season of the year and during coldest hours of the day (Buchan 1890, XXXVI).

High wind speeds are not only during storms or at gusts on individual days present, but also average wind speeds show large values (McClatchey 1996) and only few days exist on which winds are light (Ward *et al.* 1985). According to Ward (1980) wind speeds lie between 11 m/s (~24 mph) and 18 m/s (~40 mph), with maximum values reaching up to 68 m/s (~152 mph) and in 50% of the time, wind speeds are higher than 18 m/s (~40 mph). The annual mean wind speed measured between 1979 and 1986 on Cairn Gorm was 13.3 m/s (~30 mph) and the "average winter maximum monthly mean speed was 18.9 m/s (~42 mph)" (Price 2000, 157).

The west of Scotland is most exposed to winds from the Atlantic, but also in eastern Scotland, areas are close to the track of Atlantic depressions. In the beginning of a depression, winds come from south or southwest and change later to west or northwest (MetOffice 2013c). These directions are the most frequent ones in general. Northern winds bring mostly clear and cold weather (Ward 1980) and consist of dry air, southern winds are the rainiest (Buchan 1890, xl) and are often responsible for cloudy weather (Ward 1980).

Average annual rainfall is at least 1700 mm (MetOffice 2013c). Because of the maritime climate, precipitation can fall "as snow or rain at any time in the winter, with a large gradient in precipitation between the western seaboard and the relatively drier east" (Purves *et al.* 2003, 344). In winter, about 55 wet days with more than 1 mm precipitation occur over the Grampian mountains and more than 60 days in the west (MetOffice 2013c). Buchan (1890, XXIII) mentions that until the end of the 19th century, the largest annual rainfall has been measured on Ben Nevis, being 3300 mm.

Snowfalls are generally confined to the months between November and April but are in higher regions also possible in October and May (MetOffice 2013c). Ward (1980) even writes that in the Cairngorms, snow lies mostly during the whole year and that snowfall is possible all year round. Strong snowfalls are seldom in Scotland, but little snow may fall on several consecutive days enabling an accumulation of considerable depth in total.

However, the Scottish climate is not always wet, in contrast, Buchan (1890, XXIV) mentions that "the most striking feature of the climate of Ben Nevis is the repeated occurrence of excessive droughts".

A further special phenomenon in Scotland, described by Buchan (1890), is hill fog. "For the months of November, December and January this was observed for almost 80% of the time and only in May and June did the frequency fall to about 55%" (Roy 1983, 3). In winter, snow crystals can grow at objects "at an astonishing rapid rate" (Buchan 1890, XXXIX) when fogs in vapour-loaded air and deep air temperatures are drifted with high velocity over the terrain.

To sum up, Dawson (2009, 6) closes with the following sentence: "we therefore live on the edge of warm and cold, of rainfall and drought, of storminess and cold conditions." The Scottish climate is highly variable, temporally as well as spatially. The weather can change very fast and differs heavily between regions because of the proximity of coast and highlands (Purves *et al.* 2003) and because of the small scale changes in land use (Dawson 2009). Roy (1983, 2) mentions, for example, that "even in summer, were much more severe than had been realised previously" or Barton & Borthwick (1982, 228) state that "yet as much as 15 percent of mainland Scotland lies above 500 m and increasingly large numbers of people go into the uplands for active recreation at all times of the year, where they can be subjected to weather conditions far more severe than they experience in normal surroundings". Additionally, weather factors show frequently extreme values (see table 3.1) and many people have been surprised by the mercy of the Scottish weather. Especially, wind speeds are characteristic and as "wind is probably the most significant feature of the climate of the Cairngorm mountains" (McClatchey 1996, 42).

Table 3.1: Extreme climatic conditions in Scotland.

Variable	Value	Date	Location	Author
Lowest Daily Maximum Temperature	-15.9 °C	29.12.1985	Fyvie Castle	Met Office (2013)
Lowest Daily Minimun Temperature	-27.2 °C	11.02.1895	Braemar	Met Office (2013)
	-27.2 °C	10.01.1982	Braemar	Met Office (2013)
	-27.2 °C	30.12.1995	Altnaharra	Met Office (2013)
	-16.5 °C	12.01.1987	Cairn Gorm	McClatchey (1996)
	-12.6 °C	December 1886	Ben Nevis (between 1883-1887)	Buchan (1890)
Maximum Gust Speed	170 mph	20.03.1986	Cairn Gorm	McClatchey (1996)
	142 mph	12.02.1985	Fraserburgh	Dawson (2009)
	143 mph	until 1966	Cairn Gorm	McClatchey (1996)
	120 mph	February 1885	Ben Nevis (between 1883-1887)	Buchan (1890)

3.5 Scottish Avalanches in Research

Ward R., Languir E. and Beattie B. belonged to the first people who were concerned with avalanches in Scotland and their publications date around 1980. The spatial extent of their work was limited primarily to the Cairngorms and not many further publications about Scottish avalanches have been published since then (see table 3.2 for an overview).

Table 3.2: Overview of literature related to Scottish avalanches (non conclusive).

Author	Year	Title	Description	Subject
Patey, T.	1969	Post-War Winter Mountaineering in Scotland	Summary of early climbing history in Scotland, including experiences of the author, covering years until the season 1959-60.	Mountaineering
Langmuir, E.	1970	Snow Profiles in Scotland	not available	Avalanche Activity
Ward, R.	1980	Avalanche Hazard in the Cairngorm Mountains, Scotland	Summary of the avalanche hazard in the Cairngorm mountains with the aim to define different factors which influence the avalanche activtiy as basis for a forecasting model.	Avalanche Activity
Ward, R.	1981	Snow Avalanches in Scotland with Particular Reference to the Cairngorm Mountains.	not available	Avalanche Activity
Ward, R.	1984	Avalanche Prediction in Scotland: I. A Survey of Avalanche Activity	Summary of avalanche activity and description of Scottish avalanche characteristics considering various data sources.	Avalanche Activity
Ward, R.	1984	Avalanche Prediction in Scotland: II. Development of a Predictive Model	Description of meteorological conditions which favour avalanche activity, as basis for a forecast model.	Avalanche Forecasting
Ferguson, R.	1985	High Densities, Water Equivalents, and Melt Rates of Snow in the Cairngorm Mountains, Scotland	A summary of results from snowpack investigations carried out in 1984 in the western Cairngorms.	Snow
Ward, R.	1985	Snow Profiles and Avalanche Activity in the Cairngorm Mountains, Scotland	Comparison of Scottish snow profiles with those from other countries. Detailed characterisation of Scottish snow profiles is provided.	Avalanche Activity
Ward, R.	1985	Geomorphological Evidence of Avalanche Activity in Scotland	Description of widespread small-scale geomorphological features which occur at several locations in Scotland and indicate avalanche activity are provided with focus on Lochnagar.	Geomorphology
Davison, R. & Davison, S.	1987	Characteristics of Two Full-Depth Slab Avalanches on Meall Uaine, Glen Shee, Scotland	Two slab avalanches which released in Glen Shee are described in detail.	Avalanche Activity
Ballantyne, C.	1989	Avalanche Impact Landforms on Ben Nevis, Scotland	Effects of avalanches on the terrain are analysed.	Geomorphology
Luckmann, B.	1992	Debris Flows and Snow Avalanche Landforms in the Lairig Ghru, Cairngorm Mountains, Scotland	22 sites in Lairig Ghru are identified based on morphology and distinctive surface sediment sorting patterns. Landforms such as avalanche boulder tongues or roadbank tongues are described.	Geomorphology
Gordon, J.	1993	The Cairngorms	Description of the Cairngorm Massif with special concern of geologic characteristics and landforms related to glaciation.	Description of the Area
Purves, R.S. et al.	1998	The Development of a Rule-Based Spatial Model of Wind Transport and Deposition of Snow	A cell-based approach is applied on meteorological, snowpack and topographical data to model snow accumulation patterns.	Snow
Purves, R.S et al.	1998	Avalanche Forecasting Experiments for Torridon: A Feasibility Study for sportscotland	Example study for Torridon aiming to assess, if it is possible to make forecasts using data from other areas where the SAIS is currently operating.	Avalanche Forecasting
Purves, R.S. & Sanderson, M.	1998	A Methodology to Allow Avalanche Forecasting on an Information Retrieval System	SIRE, an Information Retrieval System, is applied on data from the SAIS and the SLF (Switzerland) to test its effectivity in forecasting.	Avalanche Forecasting
Barton, B. & Wright, B.	2000	A Chance in a Million? Scottish Avalanches	Many different subjects concerning Scottish avalanches are covered, including a short history, physical aspects of avalanches, the SAIS story as well as tips for survival and rescue. Many case studies are described.	Avalanche Activity
Harrison, J. et al.	2001	Climate Change and Changing Snowfall Patterns in Scotland	Socio-economic and environmental implications of climatic changes due to changes in snowfall are analysed. Scottish snowfall patterns over the latter part of the 20th century are analysed, expert opinions on effects on Scottish economy and environment are questionned in a survey and predictions for future changes are provided.	Snow
Purves, R. et al.	2003	Nearest Neighbours for Avalanche Forecasting in Scotland - Development, Verification and Optimisation of a Model	Cornice, a nearest neighbour model which supports the avalanche forecasting in Scotland is presented. Results of the model verification and testing are provided.	Avalanche Forecasting
Posdnoukhov, A. et al.	2008	Applying Machine Learning Methods to Avalanche Forecasting	Support Vector machines are applied to data from Lochaber to assess the applicability in avalanche forecasting.	Avalanche Forecasting
Moss, G.	2009	Avalanche Hazard and Visitor Numbers - a Study in Lochaber, Scotland	Visitor numbers from Ben Nevis and Aonach Mor are combined with weather data and avalanche danger levels with the aim to assess how much influence the forecasts have on the behaviour of recreationists.	Avalanche Hazard
Barraclough, T.W. et al.	2010	Snow in Scotland: Snowmicropen Analysis of Natural and Artificial Snow Samples	Samples of Scottish snowpits representing six characteristic snow types are analysed using SnowMicroPen measurements and are compared with artificial snow.	Snow
Diggins, M.	2010	Report for Winter 2009 / 10	Report of the winter season 2009/10, written by the coordinator of the SAIS, including an overview of the avalanche activity, weather situations and developments in the SAIS as an organisation.	Avalanche Activity
Floyer, J.	2010	Report: Scottish Snow and Avalanche Database	Description of the data sets from the Met Office and the SAIS.	Avalanche Activity
Sharp, B. & Whalley, D.	2010	Survey of Scottish Avalanche Incidents (1980-2009)	Overview of a 30 year avalanche accident base where the Mountain Rescue Committee has been involved.	Avalanche Activity
Diggins, M.	2011	Report for Winter 2010 / 11	Report of the winter season 2009/10, written by the coordinator of the SAIS, including an overview of the avalanche activity, weather situations and developments in the SAIS as an organisation.	Avalanche Activity
Posdnoukhov, A. et al.	2011	Spatio-Temporal Avalanche Forecasting with Support Vector Machines	A machine learning approch with the application of a Support Vector Machine for forecasting is tested on data from the are Lochaber in Scotland	Avalanche Forecasting
Diggins, M.	2012	Report for Winter 2011 / 12	Report of the winter season 2009/10, written by the coordinator of the SAIS, including an overview of the avalanche activity, weather situations and developments in the SAIS as an organisation.	Avalanche Activity

3.5.1 Publications of R.G.W. Ward in the 80s

Ward (1980) investigated until then known avalanches and wrote a survey of the avalanche activity in the Cairngorms with the aim to establish a first basis for the development of a future avalanche prediction model. Factors determining avalanche activity are found and analysed, leading to the following conditions which show highest hazards (Ward 1980, 39):

- Thaw periods
- During and after storms
- Cold periods after strong snowfall
- Persistent cold temperatures

Ward's overall conclusions include three main points: 1) "accident reports represent only a very tiny percentage of the total number of avalanches in Scotland". 2) "Avalanche danger remains the exclusive problem of the mountaineer and cross-country skier". 3) "Only two reports exist of an avalanche crossing a road in Scotland, and a train or car has never been hit" (Ward 1980, 33). Concerning forecasts, Ward (1980, 40) is still sceptic and mentions: "how big a risk depends upon the individual's character and experience, but the right to take risks is a basic mountaineering ethic, and the existence of a risk essential to the enjoyment of the sport".

Ward (1984a) is a more comprehensive survey of the avalanche activity at that time. About a thousand of avalanches which occurred over a time period of two hundred years are analysed. Several sources such as newspapers, journals, books, reports, statistics, police reports, field observations and two helicopter flights are used as data basis, but the reliability on these data is in many cases limited. Further drawbacks include that three quarters of the avalanches, of which about 60% were observed during two helicopter flights, belong to the Cairngorms and that only few data for periods with bad visibility exist, when probably high avalanche activity took place (Ward 1984a).

If possible the *location, date, time, weather circumstances, morphological avalanche type, slope, aspect, dimensions of the avalanches* and the *way of triggering* is recorded for each avalanche (Ward 1984a). Using tables and simple graphs such as histograms, the different properties of the avalanches are summarised and the following three main conclusions are found: 1) "avalanches are a common occurrence in Scotland and may be a hazard in recreation". 2) "Attempts to classify avalanches [...] to only one or two kinds of weather conditions are misleading". 3) "Some avalanches in Scotland may be significantly larger than is commonly realised" (Ward 1984a, 107).

In a third paper, Ward (1984b) summarises the steps which lead from the data to a forecasting model. Most important hereby are the meteorological conditions at the time of avalanche releases. Ward (1984b, 112) emphasises five situations, in which avalanche danger is probably highest:

- During precipitation and drifting
- When grain rounding takes place soon after deposition
- When free water is produced during thaws
- During persistent cold temperatures which cause reduced bond formations
- When depth hoar formation takes place

Based on these situations, six hypotheses are formulated and tested using cumulative percent frequency surfaces. Three different situations which lead to an increased avalanche hazard result (see table 3.3).

3.5.2 Avalanche Danger Levels and Visitor Numbers

Moss (2009) analyses visitor numbers from Ben Nevis and Aonach Mor in combination with weather data and avalanche danger levels, to assess if the forecasts play a significant role in the planning of winter tours. Weather data is transformed into a variable 'visitor enjoyment', using a scale ranging in six classes from *'clear skies / light winds / dry'* to *'storm conditions'*. Visitor numbers are used as percentage values below or above the average. Results show that by trend fewer people are in the backcountry if a high danger level is published. However, it is not possible to say, if the forecasts themselves or just the often correlated unfriendly weather conditions are the crucial factor.

Table 3.3: The three situations leading according to Ward (1984a) to an increased avalanche hazard (summarised from Ward 1984a, 129).

Situation	Avalanches are...
Fresh snowfall	likely after 200 mm of fresh snow fallen during a snowstorm (+ 50 mm on a single day) - even less snow is needed if it has been cold for several days - even less snow is needed if it has been thawing for two or three days - even less snow is needed if there has been wind during the snowstorm
Cold spells	likely after 6-7 days with maximum temperatures below -4 °C to -5 °C very likely after 14-15 days of maximum temperatures below -4 °C to -5 °C - even more likely if a very cold day (about -6 °C) or a warmer day with temperatures up to -2 °C or higher follows
Thaws	probable after 3-4 days with mean maximum temperature about 2 °C - even more likely if temperatures about 0 °C or higher follow - avalanches may not fall on the warmest day

Additionally, a tendency that weekend sportsmen make their decisions less on the basis of the forecasts than on their intended aim or plans for the day is found (Moss 2009). This correlates with the 'Consistency' heuristic mentioned by McCammon (2002) and McCammon (2004).

3.5.3 Survey of the Mountain Rescue Committee

Sharp & Whalley (2010) provide an overview of the avalanche situation in Scotland considering all avalanche accidents between 1980 and 2009, in which the Mountain Rescue Committee has been involved. The avalanche accidents account for about 2% of the total mountaineering accidents, but 7% of all fatal incidents are caused by avalanches. 158 accidents involving 431 people are registered in the considered period. 59 persons died, 231 were injured and 141 could escape without any injuries (Sharp & Whalley 2010).

The amount of avalanche accidents is highly variable. Figure 3.2 shows the number of accidents per year and Sharp & Whalley (2010) state that one can recognise a decreasing tendency, which can be explained with a stronger security culture, better preparation of the recreationists and the decrease in the frequency of nice weather.

The season for winter sports in Scotland lasts about six months, from November to April, whereas in February most accidents occur. But as only accidents with a response of the Rescue Committee are considered, the actual number of avalanche incidents is underestimated (Sharp & Whalley 2010)

The three well-known winter sport areas Ben Nevis, Glencoe and Northern Cairngorms account for 71% of the avalanche accidents (Sharp & Whalley 2010) and table 3.4 lists further areas showing high avalanche activity.

Table 3.4: Geographical areas where avalanche incidents occur (Sharp & Whalley 2010, 9).

	Number	Percentage
Ben Nevis (inc. Grey Corries and the Mamores)	45	29
N. Cairngorms (mainly northern corries)	35	22
Glencoe (inc. BEM and AE Ridge)	32	20
N. Highlands (inc. Torridon, Fannichs, Kintail, Skye)	19	12
Creag Meagaidh	12	8
S. Cairngorms (inc. Lochnagar and Angus glens)	9	6
S. Uplands/Highlands	6	3

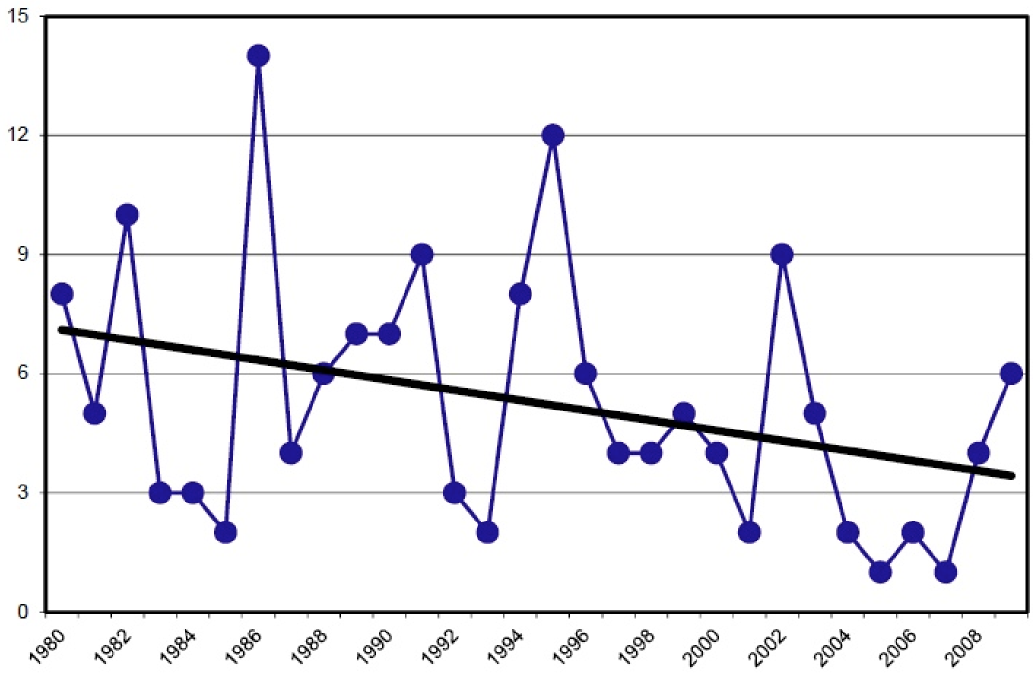


Figure 3.2: The number of avalanche incidents each year between 1980 and 2009 (Sharp & Whalley 2010, 3).

3.5.4 Latest Information

Seasonal reports that summarise most important information concerning avalanche activity and the organisation of the SAIS are published since 2010.

The winter 2009/10 was special as "predominantly arctic conditions with a northeast and easterly airflow" were present (Diggins 2010, 1). Light winds and cold air temperatures below -5 °C led to a strong temperature gradient and "produced a situation in the snowpack with grains developing into crystal types not normally seen to this extent in Scotland". Precipitation amounts were normal, but unusually large snow accumulations developed on south-west and west facing slopes. Later in that season, the avalanche activity increased due to heavy snowfalls and strong winds which formed a new compact surface layer on top of older loose grained layers. This was "probably the first time explosives have been used to safeguard infrastructures from avalanche threat in the UK" (Diggins 2010, 7). Towards April, the sun stabilised the snow cover and avalanches only occurred during storms (Diggins 2010, 1).

As normal, first winter storms reached the Highlands in October and November of the season 2010/11, but for the first time, the official operation of the SAIS started in November. After first snowfalls, cold calm conditions followed by light south westerly winds led to a very unstable snowpack on north to east facing slopes. A thaw and freeze cycle could stabilise the snow cover between mid November and December again. Refreezing, significant amounts of new snow and strong winds led to high avalanche activity in the end of December as well as in February and March. Cold and windy periods alternated with thawing conditions, in which many wet avalanches released. In the end of March warmer temperatures and sun shine led to slow settlement and stabilisation of the snow cover. Higher danger levels were then caused by storms, before the season ended in the first weeks of April (Diggins 2011b, 1).

The winter season 2011/12 started in early December, after first winter storms occurring in late October. "This was a reminder that one should be mindful of avalanche hazard from the first day of winter" (Diggins 2012, 1). A short period of stable conditions in the end of December was soon overcome by mild temperatures, very strong winds and deep thaws. Days in January were characterised by cycles of snowfall during cold conditions and warm temperatures with thaws, leading to short term natural avalanche activity. Weak layers of facets could develop during cold drifting conditions in the end of January when clear calm conditions wiled many recreationists out, resulting in high avalanche activity. In the end of March, the snow cover shrinked considerably during thawing conditions on sunny days, but the winter returned once more with a cold snowpack and short term instabilities during storm cycles. The end of the winter dates in late May when warmer conditions were present (Diggins 2012, 1).

Summarising, mainly the following conditions determined avalanche activity during the last three seasons:

- **Cold conditions over several days** with the development of weak layers increase avalanche activity.
- **Heavy snowfalls and strong winds** with drifting processes taking place increase avalanche activity.
- **First winter storms** of the season increase avalanche activity.
- **Storm cycles in spring time** increase avalanche activity.
- **Cold conditions followed by strong thaws** (sometimes combined with winds and snowfall) increase avalanche activity.
- **Thaw and freeze cycles** as well as mild conditions (higher air temperatures and stronger incoming radiation) in springtime have generally a stabilising impact on the snow cover.

4

Data and Software

The main data used in this thesis is from the SAIS, consisting of different data sets that are described in section 4.1. Additionally, weather data of the Met Office in Edinburgh and a Digital Elevation Model are used. Their properties are presented in sections 4.2 and 4.3. In appendix A.1 further details of the data sources are provided. An overview of the attributes included in the different data sets is given in table 4.1. In the second part of the chapter, short background information about the main used software is provided.

4.1 SAIS Data

The SAIS data has been collected from the beginning of the operation in 1988 until the end of the last winter season, May 2013. The majority of the data is published on the webpage *www.sais.gov.uk* and can be downloaded by everyone.

4.1.1 Reported Avalanches

The reported avalanches are stored as a point data set, in which each avalanche is represented as a point. The avalanche points are described with 19 attributes (see table 4.2 for a detailed description) and information from different data sources is combined. “SAIS observers, verbal and written reports from winter mountain activists and non mountain users” (Diggins 2010, 2) are able to report avalanches, mainly using a form on the SAIS webpage (see figure 8.14). The data is verified by the SAIS before being added to the data set (Diggins 2010, 8).

The avalanche data set can be visualised on the webpage of the SAIS as red dots on a map (see figure on title page) or as heat map. For this thesis, the data is available as an excel table that contains the attributes as columns and the avalanches as rows. The data includes numerical as well as textual values. Some example cells are shown in figure 4.1.

4.1.2 Snow Profiles

To forecast the avalanche danger level for the following day, snow profiles and field tests are carried out by the observers of the SAIS, either by ski or by foot (Diggins 2010). Information about the snow profile site, characteristics of different layers in the snow cover and weather as well as weather prognosis data are combined and should be available on a daily basis for every SAIS area.

Every morning, the observers decide on where to go for digging a snow profile that represents the most hazardous conditions of the day (Diggins 2010). The findings are reported on paper and

Table 4.1: Overview of the data sets available from the SAIS.

Data set		Nr of Attributes	Attributes	Temporal Coverage
Reported Avalanches		19	ID, Type, Fall Length, Min Width, Max Width, Min Fracture Depth, Max Fracture Depth, Altitude, Aspect, Gradient, Release, Release Attempts, Comments, Easting, Northing, Date, LonLat, Size, Region	25.12.1990 - 05.05.2013 (end of Season)
Snow Profiles Header Information	**Book 1**	19 (until 2007) 36 (since 2007)	Date, Area ID, ***Precip Code id, Observed Aval. Hazard***, Forecast Aval. Hazard, Status, Date, ***Obs, Grid, Alt, Aspect, Incline, Location***, Air Temp, Wind Dir, Wind Speed, Cloud, Drift, ***Total Snow Depth***, Foot Pen, ***Ski Pen***, Avalanche Code, ***Comments, Summit Air Temp, Summit Wind Dir, Summit Wind Speed, Rain,*** ***Max Temp Grad, Max Hardness Diff***, No Settle, Snow Index, Insolation, Crystals, Wetness, AV Cat, Snow Temp	20.12.1990 - 13.03.2013 NCs: 20.12.1990 - 13.03.2013 LO: 14.01.1991 GL: 04.01.1992 SC: - ME: 16.12.1996 ***XXX***: only since 2007
	Internet Files	34	Date, Area, Obs, Grid, Alt, Aspect, Incline, Location, Air Temp, Wind Dir, Wind Speed, Cloud, Precip Code, Drift, Total Snow Depth, Foot Pen, Ski Pen, Rain at 900, Summit Air Temp, Summit Wind Dir, Summit Wind Speed, Observed aval. hazard, Forecast aval. hazard, Avalanche Code, Max Temp Grad, Max Hardness Grad, No Settle, Snow Index, Insolation, Crystals, Wetness, AV Cat, Snow Temp, Comments	NC: 1990-2013 LO: 1991-2013 GL: 1992-2013 SC: 2007-2013 ME: 1996-2013
	Binary Files	22	Air Temp, Altitude, Region, Aspect, Aval Code, Time, Cloud, Comment, Date, Drift, Foot Pen, Forecast Aval Haz, Grid, Incline, Location, Obs, Obs Aval Haz, Precip Code, Ski Pen, Total Snow Depth, Wind Dir, Wind Speed	NC: 1989-2004 (without 1990/91, 1991/92, 1992/93, 1994/95, 2003/04) LO: 1992-2006 GL: 1990-2004 (without 1993/94, 1994/95) SC: 1996-2003 ME: 1996-2007 (see Table A.2. for detailed information)
	Cornice Wx	21	Date, Aval Code, Forecast Aval Haz, No Settle, Snow Index, Wetness, Insolation, Snow Temp (-10), Snow Temp, Snow Drift, Air Temp (-10), Air Temp, Wind Speed, Wind Dir, Cloud, Rain, Foot Pen, Av Cat, Crystals, Notes	NC: 1990-2004 LO: 1991-2004 GL: 1992-2004 SC: 1996-2004 ME: 1996-2004
	Profile Obs	22 (until 2007) 28 (since 2007)	Date, Time, Area, Obs, Grid, Alt, Aspect, Incline, Location, Air Temp, Wind Dir, Wind Speed, Cloud, Precip Code, Drift, Total Snow Depth, Foot Pen, Ski Pen, ***Rain, Summit Air Temp, Summit Wind Dir, Summit Wind Speed,*** Obs Aval Haz, Forecast Aval Haz, Aval Code, ***Max Temp Grad, Max Hardness Diff,*** Comments	NC: 1990-2003, 2007-2010 LO: 1993-2003, 2007-2010 GL: 1990-2003, 2007-2010 SC: 1996-2003, 2007-2010 ME: 1996-2006, 2007-2010 ***XXX***: only since 2007
Snow Profiles Depth Information	**Book 1**	9 4	ID, Profile ID, Depth, E, R_id, e_id, F, Comments, e, R ID, Profile ID, Temp, Depth	14.12.2007-13.03.1013 14.12.2007-13.03.1013
	Binary Files	9	Comments, Size, Form, Form1, Comment Depth, Temp, Temp Depth, Layer Depth, Wetness	NC: 1989-2004 (without 1990/91, 1991/92, 1992/93, 1994/95, 2003/04) LO: 1992-2006 GL: 1990-2004 (without 1993/94, 1994/95) SC: 1996-2003 ME: 1996-2007 (see Table A.2. for detailed information)
Forecasts			Avalanche Hazard Forecast: - Category Level, Stability Rose - Forecasted Weather Influences (Text) - Forecasted Snow Stability and Avalanche Hazard (Text) Observed Avalanche Hazard: - Observed Weather Influences (Text) - Observed Snow Stability and Avalanche Hazard (Text) Mountain Conditions (Text)	NC: 1993-2013 LO: 1993-2013 GL: 1993-2013 SC: 1996-2013 ME: 1996-2013
Annual Reports				Season 09/10 Season 10/11 Season 11/12
Blog				NC: 15.12.2006 - 2013 LO: 15.12.2006 - 2013 SC: 25.12.2006 - 2013 ME: 30.12.2006 - 2013 GL: 15.02.2009 - 2013
SIESAWS		5	Date Time Summit Air Temp (min and max) Summit Wind Dir (mean) Summit Wind Speed (mean)	Aonach mor temp.xls: 01.01.1992 - 14.03.2010 Bealach na ba temp.xls: 12.07.2001 - 30.04.2010 Cairngorm summit temp.xls: 01.01.1992 - 30.03.2010 Cairnwell temp.xls: 01.01.1994 - 30.03.2010 Glen ogle temp.xls: 01.08.1996 - 30.03.2010 Aonach mor wind.xls: 01.01.1992 - 14.03.2010 Bealach na ba wind.xls: 01.10.2001 - 12.05.2010 Cairngorm summit wind.xls: 01.01.1992 - 30.03.2010 Cairnwell wind.xls: 01.01.1994 - 30.03.2010 Glen ogle wind.xls: 01.10.1996 - 30.03.2010

back in the office, the information is typed into a program and stored as a binary file. Additionally, the profile data is put on the webpage (see figure 4.4, left).

For this thesis, several data sources which hold snow profile information are available. In each data set, different attributes are included and various time coverages are present (see table 4.1 for an overview and table 4.3 for a detailed description of the attributes).

- **Book1**[1]: is an excel file containing three sheets, one including the general attribute information (*'SnowProfiles'*, see figure 4.2), the second one including the layer information at several depths (*'ProfileLayers'*, see figure 4.3, left) and the third one including snow temperatures measured at 10 cm intervals (*'TempLayers'*, see figure 4.3, right).
- **Internet files**: are downloaded from the SAIS webpage as .csv files.
- **Binary files**: are stored in a complex and unorganised folder structure on the Floyer DVD and include one .dat file per snow profile.
- **CorniceWx**: stored on the Floyer DVD including an excel file for each area.
- **ProfileObs**: stored on the Floyer DVD including two excel files for each area (one until 2004 and one since 2007).

No data source exists, in which for every attribute in every area information since the beginning of the operation of the SAIS up until today is available. Therefore each data set mentioned above is used in this thesis and the data sets are combined to reach maximum coverages (see part B of this thesis).

For air temperature, wind direction and wind speed two different attributes are included, one measured directly by the observers at the location of the snow profile and the others, indicated with 'Summit', are measured at automatic weather stations (see section 4.2).

The depth information[2] of the snow profiles is in the two data sets *Book1* and the *Binary files* included. The attributes of the layering information included in Book1 have to be interpreted the following (Diggins 2011a, see section 2.2.2.1 for detailed explanation of the meaning):

- **ID**: id of the entry in the table.
- **Profile id**: reference to the id in the sheet *'SnowProfile'*.
- **Depth**: depth of the layer to which the data is referring, in [cm].
- **E**: size of the crystals, in [mm].
- **R**: hardness of the layer: F (Fist) / 4F (4 Fingers) / 1F (1 Finger) / P (Pencil) / K (Knife) / I (Ice).
- **R-id**: R converted to numbers: 0 = NULL / 1 = F / 2 = 4F / 3 = 1F / 4 = P / 5 = K / 6 = I.
- **e**: wetness of the layer: No Snowball / Snowball / Wet glove / Water drops / Slush.
- **e-id**: e converted to numbers: 0 = NULL / 1 = No Snowball / 2 = Snowball / 3 = Wet glove / 4 = Water drops / 5 = Slush.
- **F**: form of the crystals, consisting of a number and a letter as shown in figure 4.7.
- **Comments**: additional textual information about, for example, broken or rimed crystals and shear or Rutschblock tests.

[1] These bold data names are used in the next chapters to refer to the data sets described here.

[2] Depth information always means attributes that are recorded for several layers at different depth intervals in the snow cover.

ID	Avalanche_type	Fall_length	Avalanche_min_width	Avalanche_max_width	Avalanche_min_fracture_depth	Avalanche_max_fracture_depth	Avalanche_altitude	Avalanche_aspect	Avalanche_gradient	Avalanche_release	Avalanche_release_attempts	Comments	Easting	Northing	Date	LonLat	Avalanche_size	Region
3058	1024	150	NULL	20	NULL	NULL	950	80	NULL	2	NULL	Climber avalanched ascending to Goat T	299195	802915	14.11.2010 12:08:00	57.106313077663735 -3.664405358259793	0	0
3051	2021	50	3	4	NULL	NULL	NULL	NULL	38	6	NULL	Skier triggered 2 slab avalanches small -	300500	805800	14.11.2010 12:00:00	57.132505706817945 -3.644016094083652	0	0
3064	1021	300	NULL	NULL	NULL	NULL	1200	360	NULL	2	NULL	skier triggered - clear skies - cold large :	300055	804055	14.11.2010 11:35:00	57.11673837676151 -3.6506677812332593	0	0
3048	2	150	NULL	NULL	NULL	NULL	1100	315	40	1	NULL	Meall Nan Tarmachan, Killin. wet slides or	259150	738500	06.11.2010 14:00:00	56.51752629257988 -4.289519667315149	0	0
3067	1021	NULL	NULL	NULL	NULL	NULL	1150	25	NULL	1	NULL	Meall Nan Tarmachan, Killin.	259220	738780	06.11.2010 00:00:00	56.520061034677056 -4.288534415042765	0	0
3062	1	150	NULL	100	NULL	70	920	270	NULL	1	-9999	Great slab released after a few weeks s	298355	802755	10.04.2010 00:00:00	57.104691411090656 -3.678203018943386	0	0
3055	NULL	200	NULL	100	NULL	2	1000	360	55	4	-9999	Creise NE slopes. 2 X CT areas of debr	224005	751205	09.04.2010 00:00:00	56.619659275007784 -4.86877742769175	0	4
3046	1011	70	NULL	10	NULL	10	830	45	NULL	0	0		-9999	-9999	09.04.2010 00:00:00	NULL	0	0
3061	2	NULL	NULL	NULL	NULL	NULL	950	315	45	4	0	Creise NE slopes. 2 X CT areas of debr	224005	751205	09.04.2010 00:00:00	56.619659275007784 -4.86877742769175	0	4
3054	1011	100	5	10	NULL	NULL	1000	40	45	1	-9999	Between Coire nan Gamhna & Pinnacle Bu	244900	790000	06.04.2010 00:00:00	56.97530164453645 -4.552075681032206	0	0
										1	-9999	Jacobs area	299800	803100	05.04.2010 00:00:00	57.108106616325365 -3.6544949186488984	0	0

Figure 4.1: Example of the avalanche data set as excel file.

Date	Area	Obs	Grid	Alt	Aspect	Incline	Location	Air Temp	Wind Dir	Wind Speed	Cloud	Precip Code
3.5.2013 12:00	Glencoe	BF	217544	750	72	20	Buachaille Etive Mor	3.4		0	100	0
4.5.2013 12:40	Glencoe	PM	217543	890	35	38	Coire na Tulaich	1.1	190	15	100	2
5.5.2013 13:00	Glencoe	PM	192549	900	350	27	Stob coire Raineach	6.2	200	18	100	0
21.12.2007 10:30	Southern Cairngorms	PN/SN	157775	890	34	25	Meall Odhar	4.3	220	10	0	0
22.12.2007 11:00	Southern Cairngorms	PN/SN	162768	1000	30	25	Top of Corrie Fionn	-1	240	25	100	2
23.12.2007 11:40	Southern Cairngorms	PN/SN	186820	900	100	28	Coire Kander	-1	230	28	80	0
24.12.2007 10:30	Southern Cairngorms	PN/SN	252855	890	10	30	Central Gully Lochnagar	0.4	260	5	100	0
24.12.2007 10:30	Southern Cairngorms	PN/SN	252855	890	10	30	Central Gully	0.4	260	5	100	0
25.12.2007 10:45	Southern Cairngorms	PN/SN	162768	1000	30	25	Coire Fionn Glas Maol	-3.2	210	20	80	0
25.12.2007 14:37	Southern Cairngorms	PN/SN	162768	1000	30	25	Corrie Fionn	-3.2	210	20	80	N/A
26.12.2007 11:30	Southern Cairngorms	PN/SN	162769	1010	30	28	Coire Fionn Glas Maol	1	260	20	100	0

Drift	Total Snow Depth	Foot Pen	Ski Pen	Rain at 900	Summit Air Temp	Summit Wind Dir	Summit Wind Speed	Observed aval. hazard	Forecast aval. hazard
0	110	2		1	-0.8	93	9	Considerable	Considerable
1	100	1		1	-3.2	223	4	Considerable	Considerable
0	80	15		1	-1.2	225	18	Considerable	Moderate
0	50	0		0	-4.9	157	14	Moderate	Moderate
0	50	2		0	-4.8	148	20	Considerable	Considerable
0	30	1		0	-2.7	173	13	Considerable	Considerable
0	50	1		0	-2.5	150	17	Considerable	Considerable
0	50	1		0	0	146	12	Considerable	Considerable
0	50	1		0	-1.6	121	9	Considerable	High
0	50	1		0	-0.7	200	6	Low	Low
0	50	2		1	-2.5	253	25	Moderate	Moderate

Avalanche Code	Max Temp Grad	Max Hardness Grad	No Settle	Snow Index	Insolation	Crystals	Wetness	AV Cat	Snow Temp	Comments
0	0	3	266		20				-5.8	failures @ 8cm 25cm and 45cms
0	0	2	266		20				-5	layered slab persists
0	0	2	268		20				-0.3	
0	0	3	270		20				-4	mixed forms developing again
0	0	2	270	0	20	10	1		-5.7	Failures at 15cm and 31cm.
4400	0	2	274	4	20	10	1		-2.2	Moderate shears at 39cm
0	0	3	278		20				-1.3	nil
0	0	3	278	282	20				-0.5	New soft slab
0	0	3	280		20				-0.4	nil
0	0	4	280		20				-4.8	Icy surface layer
0	0	1	284	4	20	5	1		0	Moderate shear at 27cm.

Figure 4.2: Example of the snow profile data set (Internet files).

Profile_id	Depth	E	R_id	e_id	F	Comments	e	R
710	10	1	2	1	2a	Thin layer of surface hoar	No Snowball	4F
711	35	1	3	1	2a	TTfail Needles	No Snowball	1F
711	51	2	5	1	9e		No Snowball	K
709	3	3	1	2	7a		Snowball	F
709	5	0	5	1	9e		No Snowball	K
709	26	3	5	2	6		Snowball	K
709	34	2	4	2	6		Snowball	P
709	60	3	5	2	6		Snowball	K
712	5	2	2	1	1d	Other crystals present	No Snowball	4F
712	25	2	1	1	1d	Other crystals present	No Snowball	F
712	50	3	5	1	6b		No Snowball	K
713	43	0.5	3	1	3a		No Snowball	1F
713	23	0.5	4	1	2	WRB5	No Snowball	P
714	8	1	1	1	2a	Easy shear.	No Snowball	F
714	29	1.5	5	0	6b		NULL	K
714	102	2	4	1	6b		No Snowball	P
715	19	1	3	2	2	Mod. hand shear	Snowball	1F
716	17	0.5	4	2	2a	rounds	Snowball	P
715	40	1	2	1	2	WRB4	No Snowball	4F
715	60	3	4	0	6b		NULL	P

ID	Profile_id	Temp	Depth
72	20	-1.2	0
73	20	0	10
74	20	0	20
75	20	0	30
76	21	-0.2	0
77	21	-0.1	10
78	21	0	20
79	22	-3.9	0
80	22	-2.1	10
81	22	-1.4	20
82	22	-0.4	30
83	22	-0.5	40
84	22	-0.5	50
85	22	-0.5	60
86	22	-0.6	70
89	23	-3.1	10
90	23	-4.3	0
91	23	-1.2	20
92	23	-0.1	30
93	23	0	40

Figure 4.3: Example of snow profile depth information (Book1) (Left: sheet 'ProfileLayers' | Right: sheet 'TempLayers').

Table 4.2: Description of the attributes in the avalanche data set.,(AUTHOREN VON TABELLE 3.2: Langmuir (1970), Ward (1981), Ferguson (1985),Purves et al. *(1998a), Ward (1985), Ballantyne (1989), Luckmann (1992), Gordon (1993)*

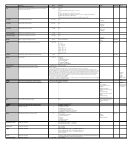


Table 4.3: Description of the attributes in the snow profile data sets.

[illegible]

4.1.3 Forecasts

To determine the avalanche hazard forecast, information of the snow profiles is combined with weather forecasts from the Met Office. In the beginning of the SAIS operation, a five-point avalanche danger scale reaching from Category 1 to Category 5 was used. Category 5 was defined as the conditions on February 6th in 1988, "when several accidents occurred and an 'all altitudes, all aspects hazard' existed" (Barton & Wright 2000, 124). The category level was only published for the observation day, because the observers did not feel confident enough to make a prediction for the forecast day as well. Nevertheless, already in the season 1990/1991 also a forecast hazard level was introduced. To avoid that only the category number is read by the recreationists without considering the text, the hazard level was written out in words. In winter 1994/95, the SAIS changed to the *European Avalanche Hazard Scale* which consists of the following five classes: 1: Low / 2: Moderate / 3: Considerable / 4: High / 5: Very High (Barton & Wright 2000).
Because of these developments, the forecasts are available in different data formats.

- **20.12.1993 to 14.04.2004**: text files can be downloaded as .zip files from the SAIS webpage.
- **14.12.2004 to 10.04.2007**: text files can be downloaded separately for each forecast.
- **13.04.2007 to 28.04.2007**: textual descriptions are published on the webpage.
- **since 16.12.2009**: a new layout is introduced (see figure 4.4, right). A stability rose is included in which the danger is given respectively to the eight cardinal directions and to two elevation bands (Floyer 2010). A hazard level table, text descriptions and glossary links are included (Diggins 2010). Download as .pdf from the webpage is possible.

4.1.4 Annual Reports

For the seasons 2009/2010, 2010/2011 and 2011/2012 an annual report summarising most important events during the winter can be downloaded from the SAIS webpage. Comments on special weather conditions or avalanche accidents are included. Example pages are shown in figure 4.5. The main points of these three reports are included in several sections of this thesis and are used as interpretation help and comparison source.

4.1.5 Blog

A blog is available on the webpage of the SAIS for each area. The observers use them to add photos to the forecasts. The SAIS has recognised "that even one picture can provide a wealth of information, additionally, if information such as field observations and experiences can be provided too then this can help the public best decide where to go." (Diggins 2011b, 9).
Example posts are shown in figure 4.6.

4.2 Weather Data

Additionally to the data of the SAIS, weather data of the Met Office in Edinburgh is available. The Met Office maintains six automatical weather stations (shown in figure 4.7, right) which belong to the SIESAWS (Severe Icing Environment Synoptic Automatic Weather Stations) (MetOffice 2013a). Three of these SIESAWS are used by the SAIS for the attributes 'Summit Air Temperature', 'Summit Wind Direction' and 'Summit Wind Speed'. Data of the station at Cairn Gorm (1245 masl) is used for the Northern Cairngorms and sometimes for Creag Meagaidh, data of the station at Aonach Mor (1130 masl) is used for Glencoe, Lochaber and sometimes Creag Meagaidh and data of the station at Cairnwell (933 masl) is used for the Southern Cairngorms.

Data of the stations Aonach Mor, Bealach Na Ba, Cairngorm Summit, Cairnwell and Glen Ogle is stored as separate files on the Floyer DVD. The .xls files containing temperature information include the attributes 'Date', 'Time', 'Maximum Air Temperature' and 'Minimum Air Temperature' at a 12 hour resolution (two measures each day at 09:00 and 21:00). The .csv files with wind information include the attributes 'Date', 'Time', 'Mean Wind Direction' and 'Mean Wind Speed' at a one-hour resolution.

4.3 Digital Elevation Model (DEM)

A DEM with 50 m resolution is used. The original data consists of 1052 .asc files (see figure 4.8, right), structured in folders according to the shortenings visible in the middle of figure 4.8. The DEM covers the whole area of Scotland, including each of the five SAIS areas without any holes (see figure 4.8, left).

4.4 Software

ArcGIS is one of the most popular commercial GIS products, developed by Esri. Different software products are integrated in various applications (esri 2013). The ArcGIS Desktop, consisting of the ArcMap and ArcCatalog, Version 10.0 is used for visualisations of the avalanche data and calculations related to the DEM.

R is an open source application language and environment which can be used for statistical and graphical data analysis. It belongs to the GNU project and has similarities to the S language. A very broad variety of applications and possibilities for extensions are ensured by R (R 2013). The R program is used for the generation of the majority of plots and statistical analysis as well as for the CA.

Excel is a table calculation program and belongs to the Microsoft Office Group. In this thesis, Excel is used to read the data, to become a first overview, to handle some pre-processing steps and to prepare the data for the import in R.

Eclipse is used to write the Java Program to read the binary files.

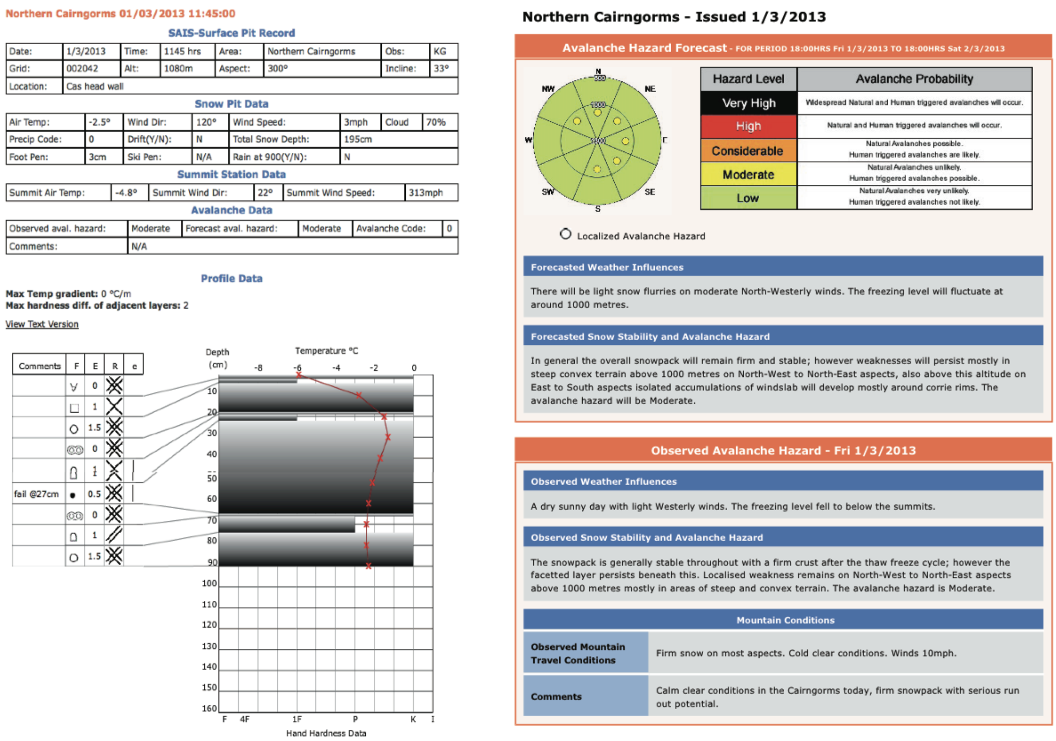


Figure 4.4: Left: example of a snow profile | Right: example of a forecast (as published on the webpage of the SAIS: (SAIS 2013b).

Figure 4.5: Example pages from the annual report of the season 2011/2012 (Diggins 2012, 1/3/4).

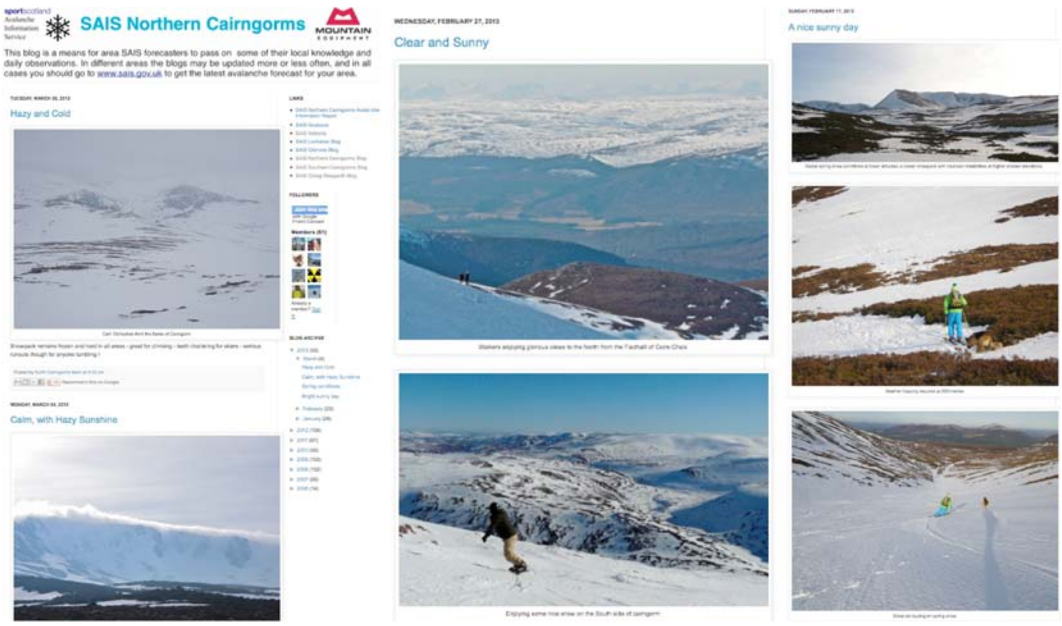
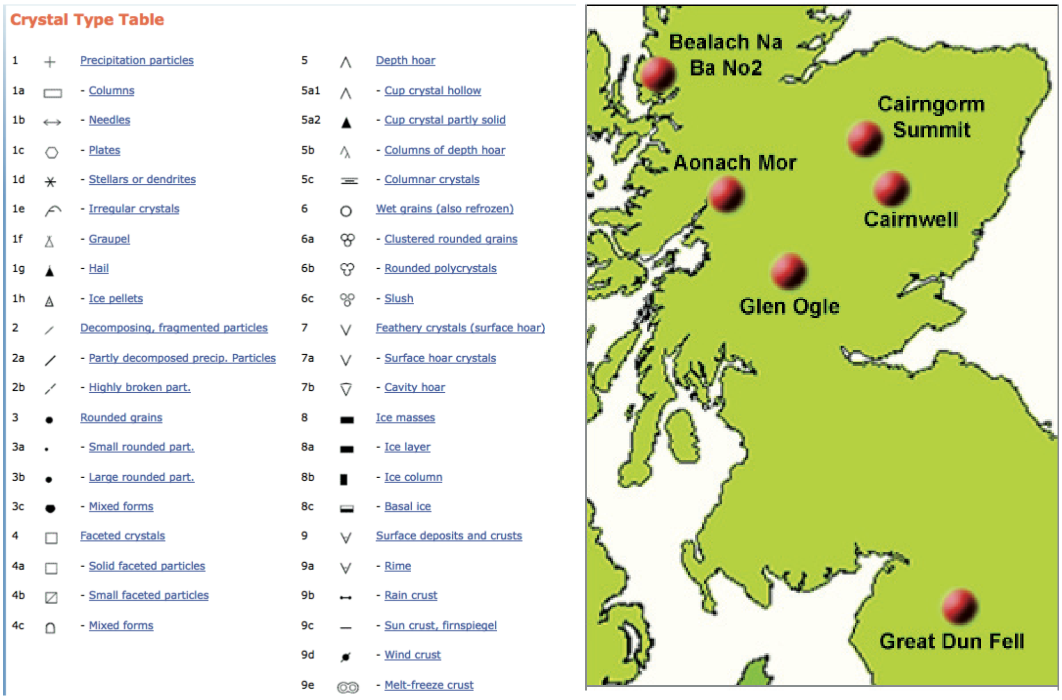


Figure 4.6: Example posts of the Blog on the SAIS webpage (SAIS 2013b: Northern Cairngorms Blog).



Crystal Type Table

1	Precipitation particles	5	Depth hoar
1a	- Columns	5a1	- Cup crystal hollow
1b	- Needles	5a2	- Cup crystal partly solid
1c	- Plates	5b	- Columns of depth hoar
1d	- Stellars or dendrites	5c	- Columnar crystals
1e	- Irregular crystals	6	Wet grains (also refrozen)
1f	- Graupel	6a	- Clustered rounded grains
1g	- Hail	6b	- Rounded polycrystals
1h	- Ice pellets	6c	- Slush
2	Decomposing, fragmented particles	7	Feathery crystals (surface hoar)
2a	- Partly decomposed precip. Particles	7a	- Surface hoar crystals
2b	- Highly broken part.	7b	- Cavity hoar
3	Rounded grains	8	Ice masses
3a	- Small rounded part.	8a	- Ice layer
3b	- Large rounded part.	8b	- Ice column
3c	- Mixed forms	8c	- Basal ice
4	Faceted crystals	9	Surface deposits and crusts
4a	- Solid faceted particles	9a	- Rime
4b	- Small faceted particles	9b	- Rain crust
4c	- Mixed forms	9c	- Sun crust, firnspiegel
		9d	- Wind crust
		9e	- Melt-freeze crust

Figure 4.7: Left: Crystal type table, describing the values included in the attribute F of the snow profile data set (Book1) | Right: SIESAWS (MetOffice 2013b).

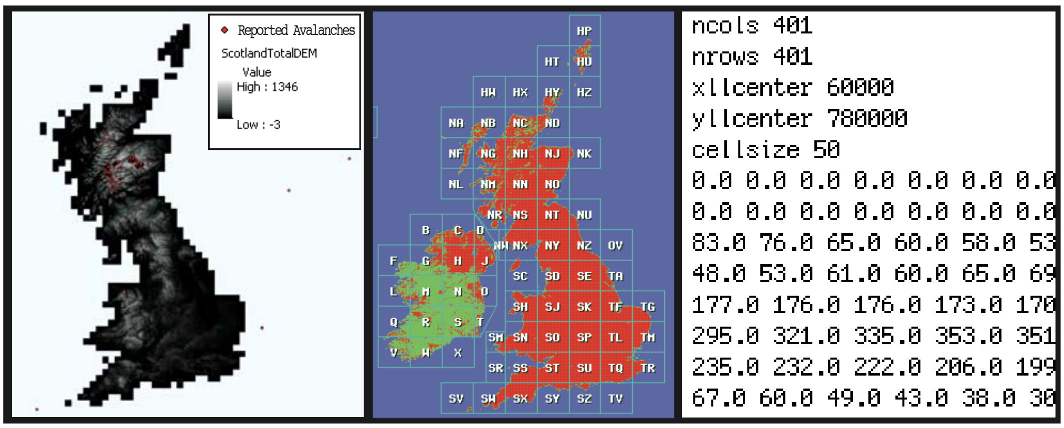


Figure 4.8: DEM for Scotland (Left: total covered area | Middle: partitions of .asc files (geograph 2013) | Right: header of .asc files.

5

RESEARCH GAP

Chapter 2 has shown that a considerable part of the current worldwide research is devoted to avalanches. A variety of studies has been conducted, pursuing different goals. Books exist which aim to describe physical properties of avalanches (e.g. Greene *et al.* 2010, Harvey *et al.* 2012, McClung & Schaerer 1993), studies in which terrain characteristics are related to avalanche activity are conducted (e.g. Guy 2011, Guy & Birkeland 2012, Guy & Birkeland 2013, Maggioni & Gruber 2002, McClung 2003, Vontobel 2011), the influences of weather variables on the release of avalanches are described (e.g. Barton & Wright 2000, Hägeli & McClung 2003, Harvey *et al.* 2012, Jomelli *et al.* 2007, McClung & Schaerer 1993, Mock & Birkeland 2000) or processes and characteristics in the snow cover related to avalanches are explained (e.g. Barton & Wright 2000, Fierz *et al.* 2009, Harvey *et al.* 2012, Johnson & Birkeland 2002, McClung & Schaerer 1993, Wiesinger & Schweizer 2001). Humans play an important role as well and have more frequently been considered or even viewed as a main factor during the last years (e.g. Adams 2005, Fredston *et al.* 1994, McCammon 2002, McCammon 2004, McCammon & Hägeli 2007, McClung 2002a, McClung 2002b, Schweizer & Lütschg 2001, Tase 2004).

Some studies have been conducted providing an overview of avalanche activity in a certain region (e.g. Casteller *et al.* 2011, Eckert *et al.* 2007, Laternser & Schneebeli 2002, McCollister *et al.* 2003, McCollister 2004, Oller *et al.* 2006, Stoffel *et al.* 1998), other authors searched for patterns in avalanche data (e.g. Birkeland *et al.* 2001, Esteban *et al.* 2005b, García *et al.* 2008, Harvey *et al.* 2002, Jomelli *et al.* 2007, Kleemayr *et al.* 2000, McCollister *et al.* 2003, Signorell 2001, Stoffel *et al.* 1998) and some services have already published avalanche patterns (e.g. Nairz 2008, SLF 2013).

Some of these authors have included suggestions for further research or improvements of their studies. Birkeland *et al.* (2001, 140) conclude that "a study integrating snow profile or stability-test data observed during avalanche extremes, and atmospheric patterns during those extremes, would also be useful". Conclusions of Sharp & Whalley (2010) include that the analysis would have been more thorough if available weather data, information about affected people and danger levels would have been considered as well and "this would be a very time consuming survey but possibly worthy of pursuing" (Sharp & Whalley 2010, 10). Signorell (2001) proposes additionally to her analysis the combination of avalanche and weather data to determine which weather factors determine avalanche releases. Hägeli & McClung (2007) propose the usage of the term 'avalanche winter regime' to enable the consideration of the avalanche activity, climatic factors and snowpack properties at the same time. Summarising, an integration of weather as well as snow cover data is proposed to analyse avalanche activity.

Chapter 3 has shown that studies considering Scottish avalanches are rather rare, but that avalanches are a persistent issue during the winter season in Scotland. Many people are out in the backcountry, and the SAIS is operating since 1988 and publishes forecasts on a daily basis for five different popular winter sport areas. Chapter 4 has shown that a lot of data has been collected, including on the one hand reported avalanches, their sizes, terrain characteristics, spatial and temporal information as well as some human aspects and on the other hand information of snow profiles about the snow cover and meteorological conditions.

This broad data base enables a comprehensive study as suggested by different authors (Birkeland *et al.* 2001, Hägeli & McClung 2007, Sharp & Whalley 2010, Signorell (2001)). Therefore the overall aim of this thesis is to give an overview of these available data sets, of the current avalanche activity in the five SAIS areas and to discover patterns in the data to improve the understanding of Scottish avalanches and to support the SAIS.

5.1 Research Questions Part A

	Overview of avalanche activity
-	What can be said about general avalanche activity in Scotland?
-	Which characteristics do the Scottish avalanches show?

As the last comprehensive study of Scottish avalanches goes back about thirty years (Ward 1984a) and many changes and developments have taken place since then, an update of Ward's study should be provided. The SAIS does not only publish forecasts but records also avalanches (see section 3.2 and chapter 4) which enables an analysis of the current avalanche activity.

The temporal and spatial distribution of the avalanches, their magnitudes and terrain characteristics of avalanching slopes will be analysed. As much information as possible should be extracted from the comments to enable insights into trigger types, related activities or the number of people involved. The results will primarily be compared to Ward's conclusions, but also to other studies to locate the Scottish situation in a broader context.

5.2 Research Questions Part B

	Combination of data sets
-	Which properties do the different data sets show?
-	Which limitations does the data have?
-	Is it possible to combine the different data sets into one comprehensive data set?
-	Is it possible to ensure satisfying data quality in the resulting data set?

Daily snow profile and weather data of the SAIS are stored in a variety of unstructured data sets (see chapter 4) and none of these is complete and reliable in every aspect. The data collected by the observers is only used for the forecast of the next day. Afterwards it remains stored without any error control or quality checks and no metadata is present. These original data sets are used in this thesis. Missing values, typing errors and inconsistencies are widely present.

Therefore the second part of this thesis consists of the cleaning and structuring of these different data sets. It will be closely looked at each attribute of every data set in every area, values will be compared and typing errors as well as inconsistencies discovered and removed. In the end, the different data sets will be combined to provide one data set that a time coverage as large as possible includes for each attribute in each SAIS area.

In literature and other studies, such data preparation steps are mentioned only seldom and the focus on data quality is mostly small, even though it should be one of the most important factors. The results of every data analysis are strongly dependent on the data and the reliability of the results are directly related to the quality of the input data (Haining 2003).

Therefore the data preparation is the central part of this thesis and done very detailed and carefully for several reasons. Due to the history of the SAIS (see section 3.2), probably more problematic issues are present in the data than in other cases. Such data preparation steps are often time consuming, tedious or even annoying and the SAIS is grateful for a reliable data set (pers. comm.: Mark Diggins: Co-ordinator of the SAIS). Therefore additional weight is put on the preparation of the data in such a way that it can be used easily by other people. The data set will be documented and described in detail. Lastly, a further data analysis which is dependent on the quality of of this data will be done in part C of this thesis.

5.3 Research Questions Part C

	Pattern analysis
-	Which patterns does the data show?
-	Are there typical factors or combinations of factors concerning topography, weather, snow cover and humans which lead to an increased avalanche activity?

The awareness of the avalanche hazard is not as widely present in Scotland as it could be and the SAIS had a hard time in gaining respect as well. The webpage has been a great step to get closer to the public, but up-to-date communication is still very important and steady effort has to be given. Therefore the idea of the SAIS has been to publish patterns which make it easier to reach target people and easier for recreationists to remember different factors related to avalanche hazards (pers. comm.: Mark Diggins: Co-ordinator of the SAIS).

Several services in other countries have already published patterns (e.g. Nairz 2008, SLF 2013), and have seen the value in such an approach. Additionally, examples of successful pattern analysis of similar data than is available for Scotland exist in literature (Birkeland *et al.* 2001, Harvey *et al.* 2002, McCollister *et al.* 2003, Signorell 2001). However, already developed patterns are related to the corresponding regions and because of the special maritime climate, fast changes in weather, extreme conditions and strong influence of the wind, they cannot be adapted directly to Scotland.

The fact that a broad data basis including information about avalanches, the snow cover and weather conditions exists, enables the discovery of patterns for Scotland as well. Different factors influencing avalanche activity will be analysed and most important ones are used to describe situations that show high avalanche hazards. A CA will be applied on the data and the days showing high avalanche activity will be grouped and assigned to typical hazardous situations.

6 METHODS

Different methods are applied to answer the research questions proposed in section 1.2 and chapter5, using the background knowledge, information of the study area and the data described in sections 2, 3 and 4. The variety of available data and aims of this thesis leads to a broad methodical approach.

In this chapter, background information about data analysis and data quality aspects relevant to this thesis is provided and the methodical steps for each of the three parts are summarised. Detailed processing steps are listed in appendix A.5.

6.1 Data Analysis - Background

Data are always an abstraction and a "capturing of the complexity of the real world in a finite representation so that digital storage is possible" (Haining 2003, 44). Data consists generally of objects whose characteristics are stored in attributes. These attributes are measured on either *nominal*, *ordinal*, *interval* or *ratio* scale and have a certain *resolution* and *accuracy*, which are influenced both by the measuring process and the purpose for which the data is collected (Haining 2003). As soon as data is used for a certain purpose, a variety of data analysis approaches and techniques can be applied.

Approaches can be divided into *inductive* and *deductive* ones. Inductive research can also be called 'bottom-up', as the work flow is directed from the specific to the general. Exactly reversed are deductive or 'top-down' approaches, in which working steps begin at the general and end at the more specific (Albers *et al.* 2006). A further distinction can be drawn between *quantitative* techniques, when numerical descriptions are used, or *qualitative* ones, when textual descriptions are used rather than numbers (Amaratunga *et al.* 2002).

Often, the beginning of a data analysis can be explorative, as stated by different authors (e.g. Fischer & Wang 2011, Fotheringham & Charlton 1994) and explained by Fotheringham *et al.* (2000, 65) the following: "it is generally helpful to look at a data set before any models are fitted or hypotheses formally tested." This led to the development of a set of techniques summarised as ***Exploratory Data Analysis*** (EDA). Good (1983) in Haining (2003, 181) defines EDA as "a collection of techniques for summarising data properties but also for detecting patterns in data, identifying unusual or interesting features in data, detecting data errors, distinguishing accidental from important features in a data set, formulating hypotheses from data". In a spatial context, ***Exploratory Spatial Data Analysis*** (ESDA) tools enhance the EDA with "additional methods that address the special queries that arise as a consequence of the spatial referencing of the data" (Haining 2003, 182).

Spatial Data Analysis (SDA) in general is special because it "represents a collection of techniques and models that explicitly use the spatial referencing associated with each data value or object that is specified within the system under study" (Haining 2003, 4). For storing spatial data, a model has to be chosen that can represent the data in an accurate way. It can be distinguished between *vector* and *raster* models and single objects can be represented either as *points*, *lines*, *areas* or *surfaces*. Additionally to their locational attributes, spatial objects can also be described with semantic ones (Burrough & McDonnell 1998).

Spatial data are not independent, as they are bounded to a certain location in space and several interactions and relationships, including proximity, connectivity and direction play a role (Miller 2010). Tobler's popular law from 1970 describes the fact, that spatially near located features tend to be more similar than features which are distant. This dependence can lead to implications in statistical analysis since independent data is needed in the majority of cases. A reduction of the degrees of freedom can result and significance tests can become problematic (Lloyd 2010).

Additionally to the three dimensions which represent the location of spatial data, the time has to be added as a fourth dimension because spatial objects can be related to a certain moment in time and change over time or even move in space and time (Haining 2003). Additionally, issues related to the extent of spatial objects such as size, shape or boundary properties have to be considered (Miller 2010).

Because of these implications, a ***GIS*** (Geographic Information Systems) is often used for SDA. GIS offer tools to handle spatial data, can be "an effective way to present information to the public" (McCollister 2006, 3), "allow professionals to better display and interpret our ever-increasing volumes of data" (McCollister 2006, 4) and can be used "in conjunction with other database methods" (McCollister 2006, 4).

A main method applied in GIS as well as in ESDA are ***visualisations***, which are claimed by several authors to be an ideal tool for parts of data analysis. Burrough (2001) mentions the strengths of visualisations for interpretation, re-evaluation of results, discovering of information or for identification and comparison. Fotheringham *et al.* (2000, 91) add that visualisations "offer not only the ability to investigate outliers and trends in multidimensional data, but also the ability to consider this in a spatial context". MacEachren *et al.* (1999, 331) state that visualisations enable "to quickly recognise common or mismatched spatial or temporal coverage in variables or develop an understanding of the potential extent, resolution, quality, cost, or other attributes of available data". Additionally, they mention that visualisations can be useful to find holes or errors in data sets. The understanding of parameter setting, techniques and their limitations as well as decisions on appropriate model representations can be simplified. Also in the context of using visualisations as part of a process and not just as presentation of the final results, Haining (2003, 189) explains that "data visualisation or scientific visualisation is concerned with the provision of many graphical views of a data set as part of an on-going process of understanding and gaining insight into the data". In general, visualisation is "an improved and effective method for organising, analysing, and communicating complex data" (McNeally 2008, 478) and "a powerful strategy for integrating high-level human intelligence and knowledge into the KDD process" (Miller 2010, 13).

KDD stands for ***Knowledge Discovery in Database*** and is motivated partly by the developments which have taken place in acquiring data. "Statistical techniques typically require a clean numeric database scientifically sampled from a large population with specific questions in mind. [...] In contrast, the empirical and synthetic data being generated and stored in many databases are noisy, non-numeric and possibly incomplete. These data are also collected in an open-ended manner without specific questions in mind or were generated as a byproduct of another activity" (Hand 1998 in Miller 2010, 10). This motivation leads to the definition of KDD as "techniques from statistics, machine learning, pattern recognition, numeric search and scientific visualisation to accommodate the new data types and data volumes being generated through information technologies" (Miller 2010, 10). KDD can also be described as being similar to a telescope or microscope: "it is a way for researchers to look at the data in different ways, discover something new, and formulate theories, models or hypothesis based on this novel information within the context of existing theory and knowledge" (Miller 2010, 17).

Fayyad *et al.* (1996, 29), MacEachren *et al.* (1999, 316) and Miller (2010, 11) describe the process of KDD in the following five steps, which are emphasised to be not at all a linear process, including interaction as well as iteration steps (MacEachren *et al.* 1999):

1. **Data selection**: decide about the data attributes and objects that should be used.
2. **Data pre-processing**: remove noise / delete duplicate data entries / deal with missing values and errors / enhance data with additional sources.
3. **Data reduction** and **projection**: decrease dimensions of the data to ensure efficient representation (called **transformation** in MacEachren *et al.* 1999).
4. **Data mining**: define the task / choose and apply algorithms to extract patterns from data sets.
5. **Interpreting** and **reporting**: interpret, evaluate, summarise and communicate the results.

Also other authors argue in favour of iterative and interactive approaches and MacEachren *et al.* (1999, 319) state that "it is only by repeated application of methods, with systematic changes in parameters, that a coherent picture is expected to emerge." Together with an explorative and iterative approach, several authors recommend an intensified interaction between GIS, statistics, geostatistics, spatial data analysis and visualisation techniques (e.g. Yuan *et al.* 2004, Fotheringham & Charlton 1994). Goodchild *et al.* (1992, 415) summarise that "general data exploratory methods would be of great value within GIS, especially in those situations where the data are of poor quality and there is a lack of genuine prior hypotheses". Additionally, they mention four cases in which statistics can strengthen GIS: in data rectification, data assessment, data sampling and in initial data exploration.

Statistical techniques can be divided into subclasses, for example in *descriptive statistics* that are used to summarise data properties and *inferential statistics* which make use of hypothesis testing methods and allow the examination of the likelihood that a statement is true or the estimation of parameters (Lloyd 2010). Generally, it can be differentiated between *univariate methods* that consider only one variable and *multivariate methods* where several variables can be taken into account and be related among each other (Huberty & Morris 1989). Backhaus *et al.* (2008) distinguish two groups of multivariate analysis methods. On the one side *structure testing methods*, such as regression analysis, variance analysis or discriminant analysis, are aimed to test primarily causal relationships of variables. Required is a a-priori knowledge of the researcher for being able to distinguish between dependent and independent variables. On the other side methods, such as component analysis, cluster analysis, neuronal networks or multidimensional scaling, belong to *structure discovering methods* which aim to discover relationships between variables where no knowledge is available beforehand.

6.2 Data Quality - Background

As data are always an abstraction, differences between the real world and the data are present in any case. Therefore every data set is characterised by different quality aspects, which are a measure of how well the data works for answering given questions (Haining 2003). Because of the complexity of spatial data, also the data quality include several components, which were described by Haining (2003, 62) the following: "a spatial data set involves the recording of attribute values and their coordinates in space and time. All three have implications for spatial data quality and errors in any one can have implications for the quality of the others." As data quality affects the results, it is important to detect and understand errors and uncertainties. According to Burrough & McDonnell (1998) errors are not bad as often is thought, but in contrast, "it can be very useful to know how errors and uncertainties occur [...] to improve our understanding of spatial patterns and processes. [...] A good understanding of errors and error propagation leads to active quality control" (Burrough & McDonnell 1998, 221).

Many different authors have tried to divide data quality into different dimensions to simplify its comprehensibility and measurability. Guptill & Morrison (1995) in Haining (2003) for example proposed seven dimensions: *data lineage*, *positional and attribute accuracy*, *completeness*, *logical consistency*, *temporal specification*, and *semantic accuracy*. In Burrough & McDonnell (1998) the aspects *currency*, *completeness*, *consistency*, *accessibility*, *accuracy* and *precision* are mentioned. Further, they concentrate on the following factors that affect the reliability of spatial data: *age of data*, *areal coverage*, *map scale resolution*, *density of observations*, *relevance*, *data format, data exchange, interoperability*, *accessibility*, *costs, copyrighting* and *numerical errors in the computer*.

To sum up, the most important dimensions including a high amount of data quality aspects concentrate on the four dimensions *accuracy*, *resolution*, *consistency* and *completeness* which are also proposed by Haining (2003). An overview of this classification is provided in figure 6.1.

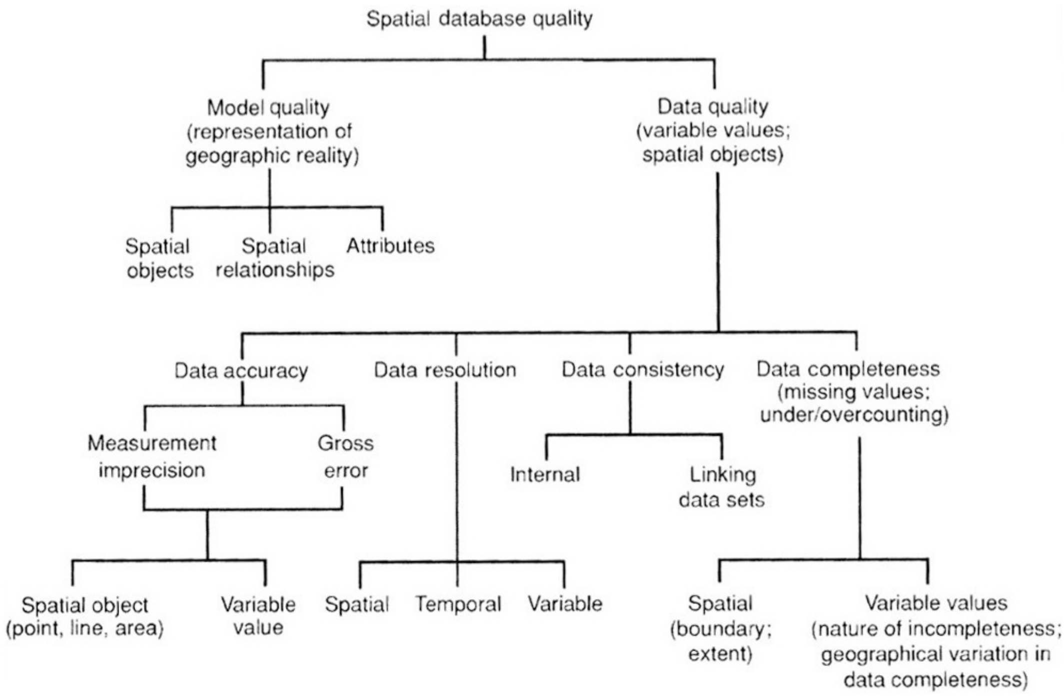


Figure 6.1: Dimensions of spatial database quality (Haining 2003, 86).

Accuracy is defined as "the difference between the value of a variable, as it appears in the database for any case, and the true value of that variable" (Haining 2003, 63), whereas the true value is not always a simple concept. It can be distinguished between *real* and *gross* errors. The latter are mostly easier detectable, for example when looking at outliers in a distribution or using descriptive statistics and boxplots (*Variable value*). For locational errors (*Spatial object*) mapping is most effective, but in some cases they only might get apparent in combination with other variables, when for example scatterplots are used. Often, background knowledge has to be applied and in case that ground truth data is available, also methods such as root mean square errors or misclassification matrices can be used to assess accuracy (Haining 2003).

Resolution is the smallest unit of measure and has an influence on several aspects of data analysis. *Spatial resolution* is the scale on which information is available for a certain area and influences the amount of represented detail. *Temporal resolution* describes the time period over which data values are aggregated and *variable resolution* is the precision at which attributes in a data set are measured. Special consideration to resolution has to be given when different data sets are merged or when scale dependent terms such as the size of a data set or the number of cases are used (Haining 2003).

Consistency is defined as "the absence of contradictions in a database" (Haining 2003, 70). Several rules exist which have to be followed in certain attributes, such as mathematical theories, formal tests, or topological rules. Some measurements have to be in a certain range because they are for example percentage values. Further, no contradictions should be included in attribute values and also consistency over time has to be ensured (*Internal consistency*). Inconsistency can be inherent in a data set or can be introduced during data analysis, for example when files are transferred, different data sets combined (*Linking data sets*) or wrong processing steps are used (Haining 2003).

Completeness has to be ensured by the model as well as by the data. Model completeness "includes whether all necessary variables are available (*Variable values*) and the spatial scale and geographic scope are sufficient (*Spatial*)" (Haining 2003, 71). Data completeness describes the amount of holes in the data, when no values and therefore no information is available for certain objects and attributes. The influence of missing data on the results of a data analysis depends on their properties, their spatial and temporal randomness and the aim of the analysis (Haining 2003).

Missing values can be dealt with in different ways. Haining (2003) explains that the easiest one is to delete all non complete entries in a data set. If the amount of missing values is small, this can be a reasonable solution. If, however, the missing values are a considerable part of the data set and are distributed randomly, a great amount of information gets lost and the significance in hypothesis testing is reduced. Therefore solutions which aim to estimate the missing values might be more appropriate. One possibility is based on imputation, where missing values are replaced, either by using the mean of the attribute for all missing values with or without considering spatial aspects, by using the empirical distribution of the attribute or by estimating missing values using regression either with or without considering random noise. An other possibility is to develop a model to estimate the missing values by using an EM algorithm (Haining 2003). Also interpolations, which estimate missing values from weighted averages of neighbouring values, or maximum likelihood estimates can be used (Goodchild *et al.* 1992).

Each of these data quality dimensions should be considered when analysing data and errors which can be prevented should not occur, but "variation of natural phenomena is not just a local noise function or inaccuracy that can be removed by collecting more data or by increasing the precision of measurement, but is often a fundamental aspect of nature" (Burrough & McDonnell 1998, 221).

Additionally to the data quality, also model quality and the interaction between data and model can affect the general quality of a study (Burrough & McDonnell 1998). Conceptualisations of the representation of space, time and attributes have to be considered (Haining 2003).

6.3 Methodical Steps in this Thesis

The authors mentioned in section 6.1 have inspired the methodical approach of this thesis. The aim is to analyse various spatial data sets with different characteristics. Therefore a variety of techniques is applied in several explorative and iterative processing steps. The five working steps of KDD are followed for answering the research questions.

The ***Data selection*** is straight forward, as the thesis aims to work with as much data as is available (see chapter 4 and appendix A.1 for detailed source information).

Within the different data sets a decision has to be made on the objects and attributes that are considered for the analysis. Mostly, as many objects as possible are used to be able to include a maximum amount of information. The selection of the attributes depends on their relevance for the specific tasks (descriptions are given in the following sections).

Data pre-processing includes a variety of working steps, at different states during the analysis, depending on the characteristics of the data. Existing attributes are changed in a new and more appropriate format, additional ones are calculated or tables are newly combined to generate useful input tables for the analysis and plotting in R. Missing values is dealt with and some attributes are enhanced using additional data sources. Some of these processes can also be seen as ***data reduction*** steps, but as abundant data is not available and the aim of this thesis is based on the total amount of available information, reduction techniques such as generalisation or aggregation are not used.

The main part of the data analysis takes place in various ***data mining*** steps using either R or ArcGIS. Many plots are generated, interpreted and served in many cases as basis for further investigation steps. Detailed descriptions of these data mining steps are given in the next sections and appendix A.5.

A summary of results from the data mining processes is ***reported*** in chapter 7 and descriptions, discussions of the most interesting aspects and ***interpretations*** are provided in chapter 8. Figure 6.2 includes an overview of the total methodical steps processed in each of the three parts.

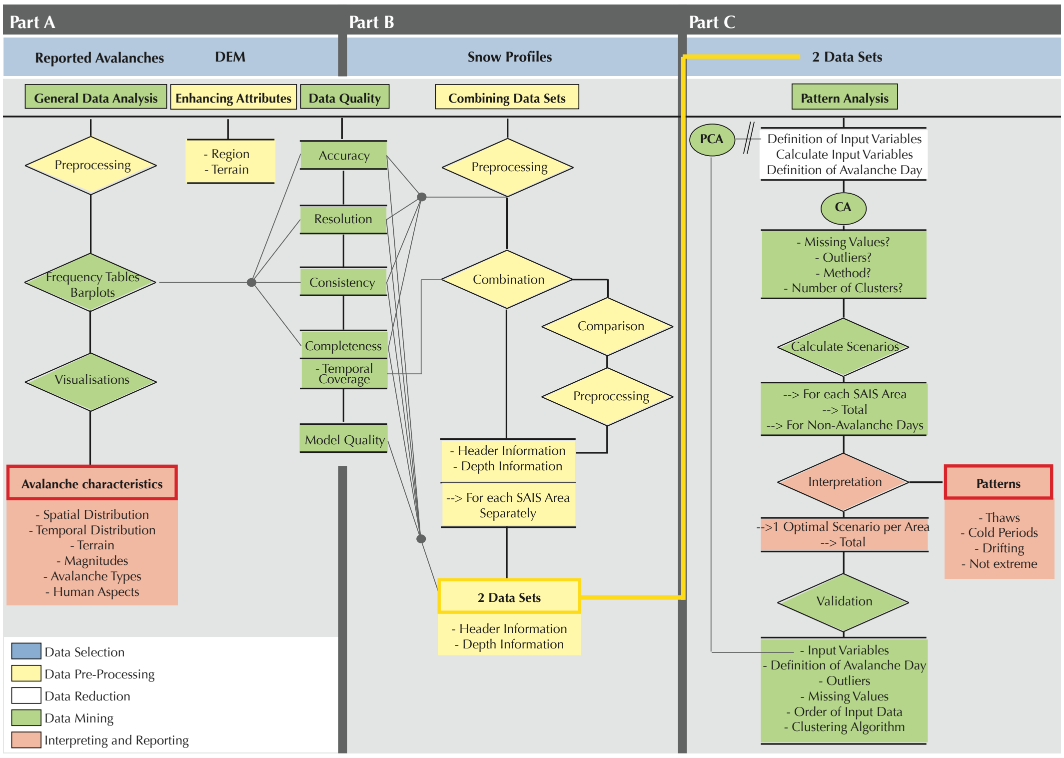


Figure 6.2: Overview of the main methodical steps processed in each of the three parts in this thesis.

6.3.1 Part A: Overview of Avalanche Activity

To give an overview of the avalanche activity, each attribute within the data set of reported avalanches is analysed in detail. Some attributes are pre-processed to enable the computing of different visualisations. Frequency tables and barplots are used to characterise avalanche properties including their distributions and typical ranges (see figure 6.3 and appendix A.5.1 for detailed processing steps). Aspects of data quality mentioned in section 6.2 and figure 6.1 are analysed, including missing values and typing errors (see appendix A.6.5 for details).

Even though missing values are an issue in the available data, general data enhancing methods such as interpolation are not applied. The original data should not be distorted, no further uncertainties introduced and there is no need to do so in order to answer the research questions. But if possible, missing values are enhanced using additional data sources.

Values of the attribute 'Region' are assigned newly to each avalanche, as they are not complete and reliable in the original data set. With plotting the avalanche points and looking at their distribution on a map, avalanches clearly belonging to one SAIS area are classified, to all other avalanches '0' is assigned. This new attribute is later used to select avalanches belonging to specific SAIS areas.

Terrain attributes including 'Altitude', 'Aspect' and 'Gradient' are extracted from the DEM. The different ascii files which produce the whole DEM for Scotland are merged and rastered in ArcCatalog. The resulting DEM is clipped to the extent of the relevant areas and the first two derivations are calculated in ArcGIS giving a gradient and aspect surface. For each surface as well as the altitude, the values at avalanche data points are extracted and stored as additional attributes in the data set.

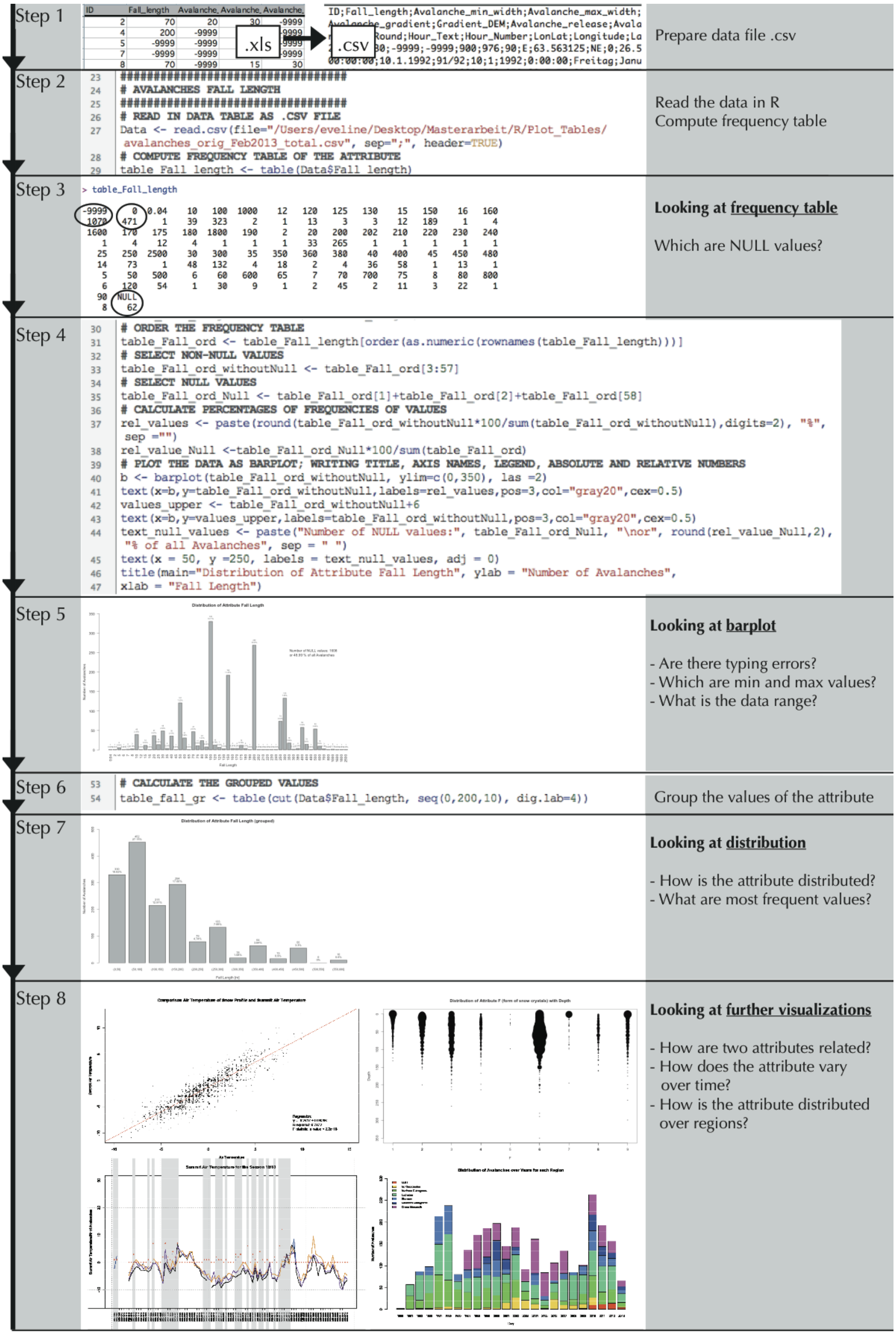


Figure 6.3: Overview and examples of the data analysis steps applied in Part A.

6.3.2 Part B: Combination of Data Sets

To establish a data set with a high data quality that includes snow profile information, the five data sets have to be combined. Because of their different characteristics, each data set is processed separately (see appendix A.5.2 for details). A close look is taken at every attribute (similarly as shown in figure 6.3 for the avalanche data set) to get a first impression of the data. The focus of the further analysis is on data quality with the temporal coverage as one of the most important aspects.

The different data sets are combined in order to maximise the temporal coverage for each attribute. Because data characteristics are not in any case identical for the different SAIS areas, all steps are processed independently for each area. The header information, including general properties of the snow cover as well as meteorological variables, and the depth information, including data for each layer, are processed separately since different steps are necessary.

- **Header Information**
  The *internet files* are used as basic data because it is thought that these published data are the official information and should therefore be reliable, complete and of good quality. Missing values in this data set are enhanced as far as possible with other data sources. The combination of the attributes 'Region' and 'Date' allows a clear and reliable indexing. The adding is done in Excel using formulas such as *INDEX(MATCH())*. To ensure consistency between the different data sources, further data processing steps are done (see appendix A.6.1). With having columns of the same attribute from different data sources in one file, the values are compared and tested on identity either directly in Excel or using R.
  Especially important is the comparison of different summit station data sets. More than 15 years before 2007 have to be enhanced using an additional data source and these attributes play an important role in part C of this thesis. For the SAIS areas Lochaber, Glencoe, Southern Cairngorms and Creag Meagaidh, information is available in the *internet files* since 2007 and in the *SIESAWS* data for the years up until 2010. For the Northern Cairngorms, data can additionally be downloaded directly from the webpage of Cairn Gorm Automatic Weather Station (cairn gorm 2013). To support the choice of the data sources which are used to enhance the data of the *internet files*, values are compared for the Northern Cairngorms in detail. Comparisons of *internet files* and *Cairn Gorm station* data are possible between 2007 and 2013, those of *internet files* and *SIESAWS* data between 2007 and 2010. Table 6.1 shows an overview of these data sets and the resolution of the included attributes. In the last three columns, all data pairs which are compared to each other are signed with a tick.
  In a last step, the columns of the *internet files* and one additional data source are combined using formulas such as *IF(ISNUMBER(Internet data);Internet data;IF(ISNUMBER(Profile Obs);ProfileObs;"NULL")*. If for one profile two values are available from both data sources, the value of the *internet files* is used.
- **Depth Information**
  The depth information is a combination of the two data sources *Book1* (sheets '*ProfileLayers*' and '*TempLayers*') and *binary files*. Until 2007, data of the *binary files* are used, after 2007 those of *Book1*. To ensure a clear indexing, each profile has a profile ID and each layer in the profile a layer ID with numbers from 1 (uppermost layer) to 20 (bottommost layer) (see appendix A.6.2 for detailed processing steps).

The resulting data sets are stored in two excel files, one including the header and one the depth information. To enable further applications of the data by other persons, each attribute is described in detail in a metadata file. Additionally, resulting graphs and plots of the general data analysis are stored in an ordered and logical way on a DVD which is available for the SAIS (see appendix A.7 for content).

The data quality of the resulting data sets is assessed similarly to the reported avalanches (see section 6.3.1 and appendix A.6.5). Limitations such as typing errors are removed and inconsistencies are overcome when ever possible.

The combination of the different data sets and the assessment of data quality is often done simultaneously as they are closely related. For better readability the results of the processing steps are summarised not chronologically but according to their content in sections 7.2.1 and 7.2.2.

Table 6.1: Different sources of summit station data.

Data Source	Attributes	Values	Unit	Area / Location	Time Span	Additionally Calculated	SP	FL	CG
Snow Profiles (SP) **(Internet files)**	Summit Air Temperature	Every day Value at midday	[°C]	NC LO GL SC ME	2007 - 2013	-		✔	✔
	Summit Wind Direction	Every day Mean of values over the last 24 hours	[°]			-		✔	✔
	Summit Wind Speed	Every day Mean of values over the last 24 hours	[mph]			-		✔	✔
Station Data (FL) **(Floyer DVD)**	Maximum Air Temperature	at 9:00 and 21:00 Mean of values over the last 12 hours	[°C]	Aonach Mor Bealach Na Ba Cairn Gorm Cairnwell Glen Ogle	1992 - 2010 2001 - 2010 1992 - 2010 1994 - 2010 1996 - 2010	Mean per day 21:00 - 21:00	✔ 09,21, Mean		✔
	Minimum Air Temperature	at 9:00 and 21:00 Mean of values over the last 12 hours	[°C]			Mean per day 21:00 - 21:00	✔ 09,21, Mean		✔
	Mean Wind Direction	Every round hour Mean of values over the last hour	[°]			Mean per day 21:00 - 21:00	✔ 09,21, Mean		✔
	Mean Wind Speed	Every round hour Mean of values over the last hour	[mph]			Mean per day 21:00 - 21:00	✔ 09,21, Mean		✔
Cairn Gorm Station Data (CG) **(Webpage Cairn Gorm)**	Temperature 1	Every half hour, measured 3min Minimum value measured in the period	[°C]	Cairn Gorm	1990 - 2013 without 1996, 2003	Mean per day 24:00 - 24:00	✔	✔	
	Temperature 2	Every half hour, measured 3min Minimum value measured in the period	[°C]			Mean per day 24:00 - 24:00	✔	✔	
	Minimum Temperature additionally calculated)		[°C]			Min of day 24:00 - 24:00	✔	✔	
	Maximum Temperature (additionally calculated)		[°C]			Max of day 24:00 - 24:00	✔	✔	
	Wind Direction	Every half hour, measured 2.5min, Every 2.5 sec Mean value measured in the period	[°]			Mean per day 24:00 - 24:00	✔	✔	
	Stdv in Wind Direction	Every half hour, measured 2.5min, Every 2.5 sec Standard deviation of the values measured in the period	[°]			Mean per day 24:00 - 24:00	✔	✔	
	Mean Wind Speed	Every half hour, measured 2.5min, Every 2.5 sec Mean of values	[m/s] until 1995 [mph] after 1995			Mean per day 24:00 - 24:00	✔	✔	
	Maximum Gust	Every half hour, measured 2.5min, Every 2.5 sec Maximum value measured in the period	[m/s] until 1995 [mph] after 1995			Mean per day 24:00 - 24:00	✔	✔	
	Minimum Wind Speed	Every half hour, measured 2.5min, Every 2.5 sec Minimum value measured in the period	[m/s] until 1995 [mph] after 1995			Mean per day 24:00 - 24:00	✔	✔	

6.3.3 Part C: Pattern Analysis

Pattern discovery in the data is done using a PCA (Principal Component Analysis) and CA (Cluster Analysis). The resulting data sets of Part B are used as basis. These are no longer seen as special snow profile information, but as information describing the situation at the corresponding date, including aspects of the weather and the snow cover.

A large data set with a high number of variables tends to become complex and unclear. A high probability of correlation and redundant information exists. Techniques to reduce such complexity include PCA. The central idea of a PCA "is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set." (Jolliffe 2002, 1). A PCA builds groups of variables (called components), which are highly correlated (Backhaus *et al.* 2008) and ordered, so that the first few components retain most of the variation present in the original data (Jolliffe 2002). A PCA consists of a variety of steps which are shown in detail in appendix A.6.6.

Originally, the idea for the application of the PCA in this thesis has been to help defining input variables for the CA (compare for example Esteban *et al.* 2005b or Mock & Kay 1992). But as most PCA algorithms can only cope with complete cases, the remaining input data would have been too small to allow a sensible application (see section 7.3.3 for detailed information). Some algorithms (for example NIPALS: Grung & Manne 1998) estimate missing values, but as already mentioned, additional estimates would have gone beyond the scope of this thesis. In the end, the input variables for the CA are decided based on several other aspects. Each attribute available in the data is analysed to decide if it is relevant or not, considering the influence of different factors on general avalanche activity (see section 2.1.2). Further, it is looked at existing studies with similar aims to see which variables were included (see section 2.3.2). Lastly, conditions which are already known as hazardous situations in Scotland are considered (see sections 3.5.1 and 3.5.4). The PCA is then applied on these input variables to see how much of the variance included in the data can be explained as a measure for the quality of the CA.

The aim of a CA is to divide not variables but events or objects into different groups. In this the current analysis, the avalanche days are grouped into different situations which can be summarised as patterns. The groups have to be built in such a way that the data in one group is as homogenous and the different groups as heterogeneous as possible (Mooi & Sarstedt 2011). Appendix A.6.7 summarises most important processing steps in detail. A variety of decisions has to be made within a CA. As they are dependent on the data and mostly made during data analysis, a detailed description is provided in the results in section 7.3.2. Here is a short summary given[1]:

1: Eleven **variables** are used for the cluster analysis, including information about air temperature, wind speed, amount of new snow and snow cover stability.
1: The threshold for an **avalanche day** is defined separately for each area based on the comparison of different results: NC: >=3 avalanches/day, LO: >=3 avalanches/day, GL: >=3 avalanches/day, SC: >=2 avalanches/day, ME: >=3 avalanches/day, Total: >=2 avalanches/day.
1: **Missing values** are not replaced. Days with at least one missing value are not omitted, but to ensure a smooth run in R, days with only one to four values are not considered in the analysis.
1: **Outliers** are detected by looking at the dendrogram of a single linkage hierarchic clustering. But as extreme outlier values are typical for Scottish conditions, they are included in the analysis.
2: As **similarity measure** the *Euclidean Distance* is used.
3: A **hierarchical clustering** algorithm with **Ward's method** is applied.
4: The **number of clusters** is defined separately for each area: NC: 5 clusters, LO: 4 clusters, GL: 4 clusters, SC: 5 clusters, ME: 6 clusters, Total: 5 clusters. The decision is based on screeplots and the comparison of results which show different numbers of clusters.
5: For the **interpretation** of the cluster solution mainly boxplots and cluster means are used.
6: The results are **validated** by comparing scenarios with different thresholds for an avalanche day, with and without outliers or omitting days, with different orders of input data and by using a different clustering algorithm.

For each of the five SAIS areas as well as for the total amount of data and non-avalanche days, various scenarios are conducted, including different thresholds for an avalanche day and several numbers of clusters. In the end, one optimal scenario for each area is selected and described in detail in section 7.3.2.4.

6.4 Comments on the Methodical Approach

With the application of SIRE has an Information Retrieval System be applied on the Scottish avalanche data (Purves & Sanderson 1998), with the development of Cornice, a nearest neighbour approach (Purves *et al.* 2003) and also a Support Vector Machine has been tested on data of Lochaber (Posdnoukhov *et al.* 2011). The aim of this thesis is to give new insights into the data and therefore new methods are used. In the literature some successful examples are described for both, applications of PCA and CA on avalanche data. Further, the method has to be combined with those used in Parts A and B of this thesis and must be applicable to the available data. Therefore characteristics such as *explorative*, *structure discovering*, *comprehensible* and *replicable* are important. At last, the method should be appropriate for the scope of a Master thesis and be reasonable for one out of three parts.

[1]The numbers in the list reference the processing step of the CA in appendix A.6.7

7

RESULTS

The outcomes which resulted from the application of the methods summarised in chapter 6 to the data presented in chapter 4 are presented next. Aspects which were analysed by Ward (1984a) are updated to the current state of the data and further insights are provided, as far as enabled by the available data.

The first part of the chapter contains an overview of the properties of the avalanches and avalanche activity in general. The second part summarises issues concerning the analysis of the snow profile data sets including their data quality and the resulting comprehensive data set is presented. The last part of the chapter describes the results of the pattern analysis from the CA. A discussion of these results and comparisons to other studies are provided in chapter 8.

7.1 Part A: Overview of Avalanche Activity

7.1.1 General Data Analysis

The analysis of the reported avalanche data set shows the following results:

Spatial Distribution
The distributions of the avalanche densities provided in figure 7.1 show no strong and steady increase in the number of avalanches, but the area they cover is getting bigger, especially in the last few years. The five ares covered by the SAIS can be distinguished clearly by the clustering of the avalanche points and the temporal development of the SAIS areas is reflected. In the season 1990/91, only the two areas Northern Cairngorms and Lochaber were considered and the areas Southern Cairngorms and Creag Meagaidh were added in the season 1996/97. Since the season 2010/11 markedly more avalanches between or outside the SAIS areas are reported, including regions such as the Northern Highlands.

The Northern Cairngorms and Creag Meagaidh show largest total number of avalanches per area (see table 7.1). The avalanches in Creag Meagaidh all have been reported after 1996 which results in a high number of avalanches per year.

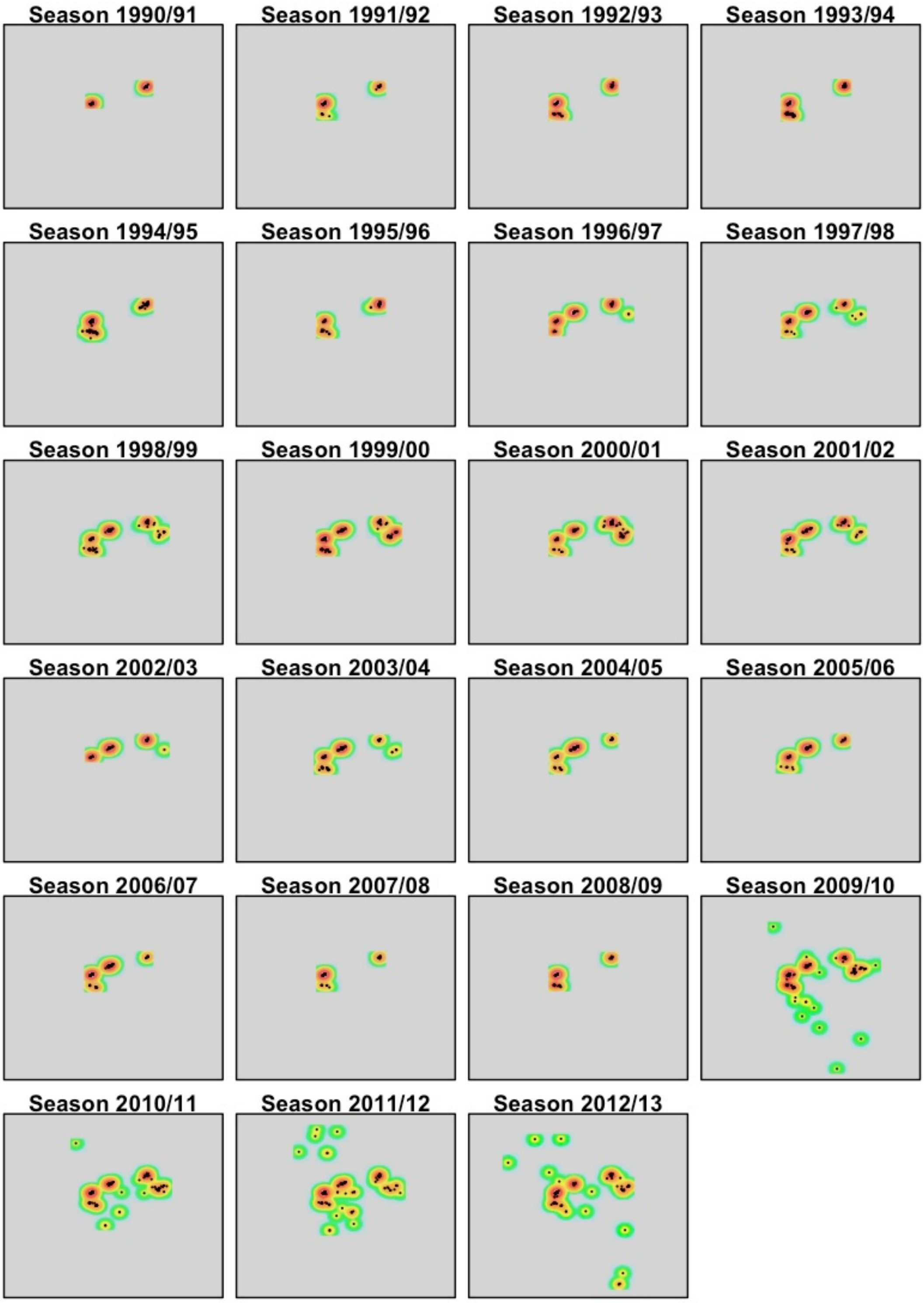


Figure 7.1: Densities of the reported avalanches for each season.

Table 7.1: Number of reported avalanches per area and season. (Few avalanches are already seasons before 2009 reported, but since 2009/2010, the number has considerably increased).*

Area	Beginning of the SAIS Operation	Total Number of Avalanches	Mean Number of Avalanches per Season
Northern Cairngorms	1988/89	720	28.8
Lochaber	1989/90	976	40.7
Glencoe	1988/89	469	18.8
Southern Cairngorms	1996/97	189	11.1
Creag Meagaidh	1996/97	689	40.5
outside SAIS	(2009/10)*	58	14.5

Temporal Distribution

The temporal distribution of the recorded avalanches can be analysed on different scales. The smallest ones are the yearly and seasonal scale which are shown in figure 7.2. First avalanches were reported in 1990, when the SAIS started operating. This leads to a 24-year or 23-season data series until the current year 2013, including a total amount of 3282 avalanches. In average, approximately 136 avalanches per year and 142 avalanches per season have occurred. Most avalanches are reported in the season 2009/2010 (261), respectively in the year 2010 (263). Also the years 1994 and 1995 show a high number of avalanches with values above 200. A number of about 80 avalanches can be seen as minimum amount of avalanches being reported per year or season.

The distribution of the avalanches throughout the months (see figure 7.3) shows that the Scottish avalanche season lasts more than half a year, from November until May, whereas in November and May less than 1% of the avalanches occur. The distribution seems to be fairly normal with highest values and approximately one third of all avalanches occurring in February and only slightly less in January and March. The avalanches are distributed more or less even throughout the week, but the frequencies are highest on weekends. In average, 540 avalanches occurred each on both weekend days and 440 on each weekday. Avalanches occur mainly between 10 am and 2 pm with clearly the highest value at 12 am. Least avalanches release in the evening between 7 pm and midnight. In the second part of the night, after midnight until 8 am, about 14% of the avalanches are reported.

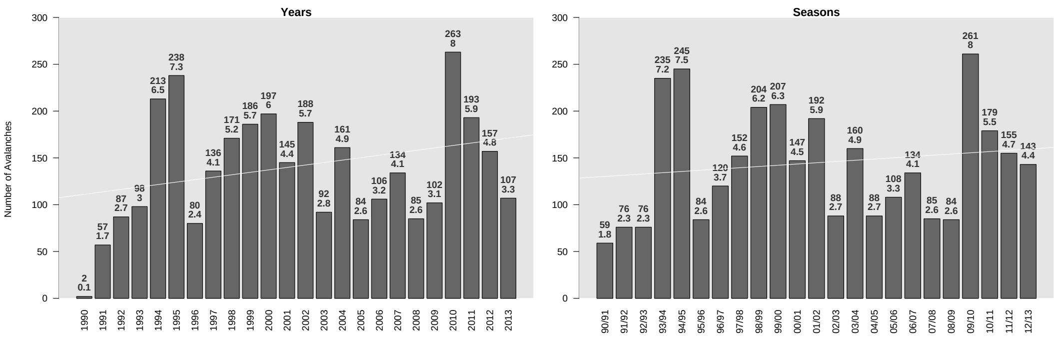


Figure 7.2: Distribution of the reported avalanches over the years 1990-2013 and the seasons 1990/91-2012/13. (Upper values: [abs], Lower values: [%]).

A detailed overview of the number of reported avalanches on a daily resolution is provided in appendix A.3.

Terrain Characteristics

Terrain characteristics of the slopes on which avalanches released include altitude, aspect and gradient (see figure 7.4). The majority of avalanches (about 56%) occurs at around 1000 (+- 100) masl and 75% release above 900 masl. Some few avalanches (less than 1%) occur below 500 masl and another small part above 1200 masl (less than 2%).

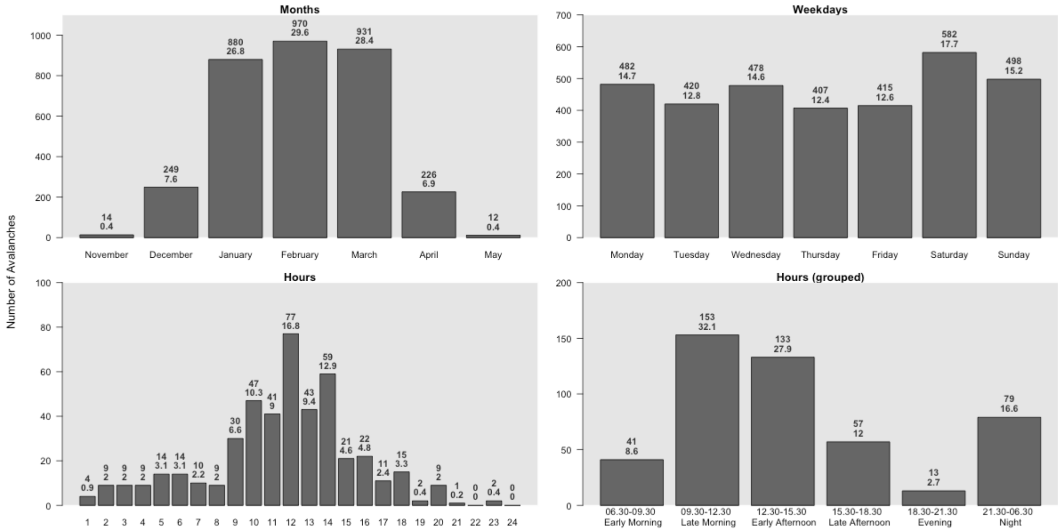


Figure 7.3: Distribution of the reported avalanches over months, weekdays and hours. (Upper values: [abs], Lower values: [%])

The distribution of the aspect shows that the majority of avalanches (about 60%) occurs on slopes with a northeastern or eastern aspect and some with a northwestern or northern aspect (nearly 30%). Only about one tenth of all avalanches release on southern or western aspects.

For the gradient, in the original data set only values for 153 avalanches are available, corresponding to 4.6% of the total amount of data. Almost half of these avalanches occurs at slopes with a gradient of 40° or 45°. About one third of the avalanches releases on slopes between 30° and 40° and 12 avalanches are recorded on slopes steeper than 45° with the highest value of one avalanche at 65°. Special within this attribute is the fact that all values are from avalanches which released after 14.11.2010. No information is available in the older data.

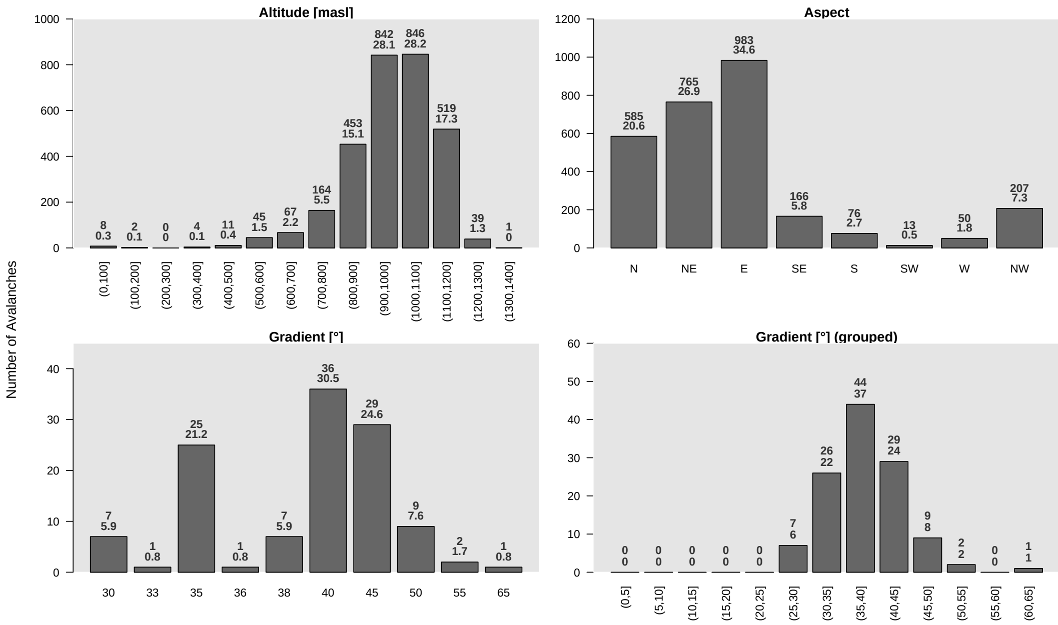


Figure 7.4: Distribution of the reported avalanches over altitudes (upper left), aspects (upper right) and gradients (original values lower left, grouped values lower right). (Upper values: [abs], Lower values: [%])

Magnitudes

Different attributes in the avalanche data set describe the magnitude of avalanches. The attribute 'Size' includes values between 1 and 5, with 5 being a very big avalanche which can destroy a whole village. The majority of the reported avalanches is relatively small, with a size of 1 or 2. Some few avalanches are reported with size 3 and two avalanches with size 4 (see figure 7.5).

The majority of avalanches travels less than one kilometre, almost half of the avalanches even less than 100 m (see figure 7.5). But few avalanches are recorded to have reached far higher fall lengths by travelling 1.6 km, 1.8 km and 2.5 km.

The fracture depth ranges from less than one centimetre to about three meters (see figure 7.6). One measure for the minimum fracture depth and one for the maximum fracture depth is included in the data set. The minimum fracture depth ranges from 10 cm to 2 m, the maximum fracture depth from 3 mm to 3 m. The majority of avalanches shows a maximum fracture depth of less than 1 m as only about 6% of all avalanches fractured deeper than one meter. Values larger than 2 m are for only 5 reported (1 avalanche with 2.5 m and 4 avalanches with 3 m of maximum fracture depth).

Also for the width two measures for the minimum and maximum values are available ranging from almost 0 to 1000 meters (see figure 7.7). The majority of the avalanches (about 80%) shows widths of less than 50 m and about 91% of the avalanches are less than 100 m wide. There are some few extreme values with one avalanche reported as having a maximum width of 1000 m, one with 750 m, one with 700 m, two with 600 m and seven with 500 m.

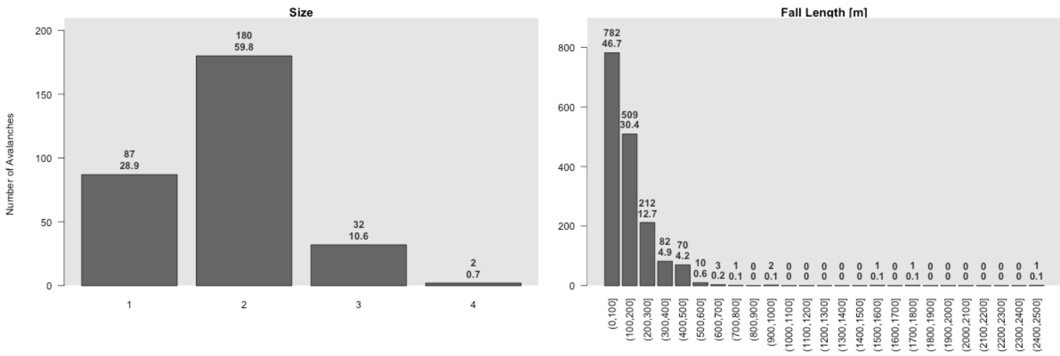


Figure 7.5: Distribution of the reported avalanches considering the size (left) and fall length (right). (Upper values: [abs], Lower values: [%])

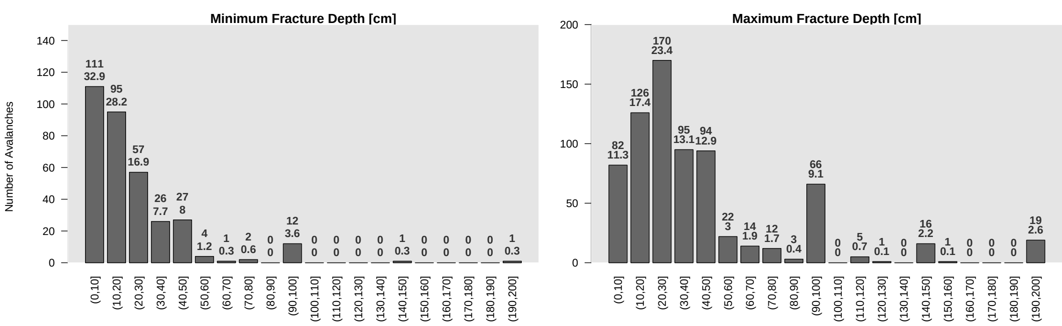


Figure 7.6: Distribution of the reported avalanches considering minimum (left) and maximum (right) fracture depth. (Upper values: [abs], Lower values: [%])

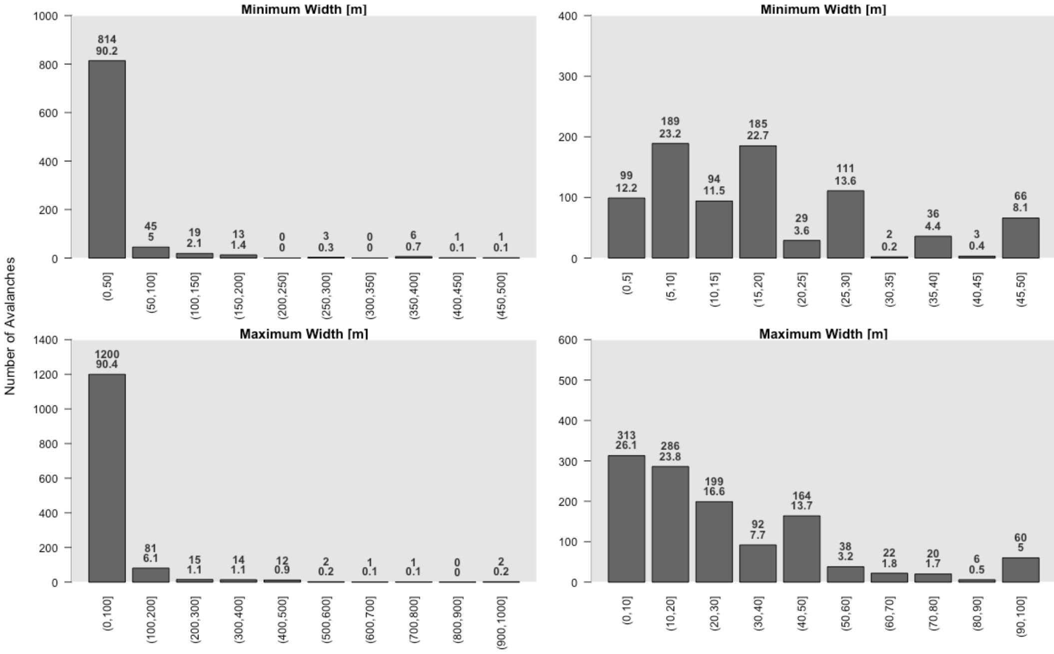


Figure 7.7: Distribution of the reported avalanches considering minimum (left) and maximum (right) width. (Upper values: [abs], Lower values: [%])

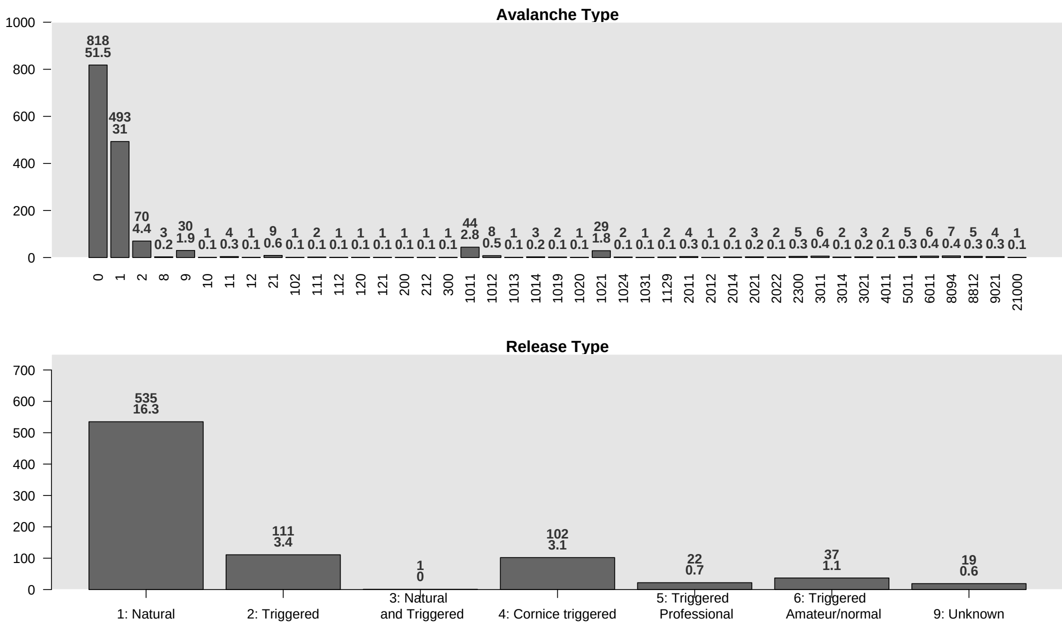


Figure 7.8: Distribution of the reported avalanches considering the avalanche type and release type. (Upper values: [abs], Lower values: [%])

Comments
The attribute 'Comments' in combination with 'Avalanche Release' and 'Avalanche Type' (see figure 7.8) provides additional information for more than 90% of all avalanches. Information about wetness, the fracture type of the avalanches, the trigger, number of people involved, related activity and injuries or fatalities are extracted.

Table 7.2: Summary of information extracted from the comments. (The numbers show the amount of avalanches that are mentioned in relation to the corresponding information).

Information	Values				
Wetness	**dry**	**wet**	**dry/wet**		
	936	518	5		
Fracture Type	**slab**	**sluff**	**slab/sluff**	**loose snow**	
	882	53	14	190	
Trigger	**natural**	**human**	**cornice**	**animals**	**unsure**
	792	297	436	3	19
Nr of People	**zero**	**one**	**two**	**three**	**four**
	3	59	97	27	11
	five	**six**	**seven**	**eight**	**nine**
	2	7	4	1	2
	ten	**group**			
	1	53			
Activity	**climbing**	**walking**	**climbing/walking**	**skiing**	**boarding**
	179	23	3	94	14
	walking/skiing	**skiing/boarding**	**observer/SAIS**	**observer/skiing**	**instructor**
	1	7	130	2	11
	ski patrol	**piste machine**	**quad**	**roadside**	
	30	2	1	58	
Outcome	**buried**	**injured**	**fatal**		
	58 people in 30 avalanches	67 people in 37 avalanches	25 people in 13 avalanches		

7.1.1.1 Summary of Avalanche Characteristics

The main points of the last section characterising the avalanches included in the reported avalanche data set are summarised the following:

Spatial distribution:

- The **locations** of the avalanches reflect the five operational areas of the SAIS.
- The distribution gets looser and the reported avalanches are not anymore bounded so strongly to the five areas since 2010.
- Most avalanches are recorded in Lochaber, least in Glencoe and the Southern Cairngorms. Highest frequencies per season show Lochaber and Creag Meagaidh.

Temporal distribution:

- The **yearly** or **seasonal** distribution show no strong general trend. During the season 2009/2010 or the year 2010 were with 261 and 263 most avalanches recorded.
- **Months** showing avalanche activity reach from November to May with highest values and in total about 85% of the avalanches occurring between January and March.
- The distribution during the **week** is balanced. On weekends approximately three percent more avalanches are recorded than during the week.
- Most avalanches release in the **hours** between 9 am and 3 pm with a clear maximum at noon. About 17% release during the night.

Terrain:

- More than half of the avalanches release at an **altitude** around 1000 masl. Less than one percent releases below 500 masl and less than two percent above 1200 masl with a maximum of 1344 masl.
- About 60% of all avalanches release on slopes which have either **aspect** NE or E. Only about one tenth of the avalanche release on southeastern to western slopes.
- Almost half of the avalanches release on slopes with a **gradient** of 40° or 45°. Values range from 30° to 65°.

Magnitudes:

- Most avalanches are classified with a small **size**, having a value 1 or 2. 10% of the avalanches are classified in size 3 and two avalanches with size 4 were reported.
- The majority of avalanches shows **fall lengths** of less than one kilometre, almost half of the avalanches even less than 100 m. Maximum values are 1.6 km, 1.8 km and 2.5 km, each reported for one avalanche.
- **Fracture depth** ranges until about three meters but the more than 90% of all avalanches have a fracture depth of less than one meter.
- **Widths** range from almost zero to 1000 m whereas about 80% of the avalanches are less than 50 m and about 91% less than 100 m wide.

Avalanche types:

- Differencing the avalanches according to their **wetness**, almost twice as many avalanches are classified as dry (936) than wet avalanches (518).
- With 882 avalanches categorised as slab avalanche and 53 as sluff avalanche, the **fracture** is much more often of a slab type.
- 792 avalanches are said to be **triggered** naturally, 297 by humans and 436 are mentioned to be related to cornices.

Human aspects:

- Only for a small amount of avalanches information about the number of people involved is available. **Group sizes** reach from 0 to 10 people, with maximum values between 1 to 3 people (59, 97 and 27 avalanches). In 53 cases a group is mentioned without information about the exact number of members.
- Different **activities** are mentioned which are carried out during the avalanche release, with a maximum value for climbing (179). Also skiing and the reporting by observers have with 94 and 130 high values. Walking, boarding and ski patrol or instructors are less often mentioned (23, 14, 30, 11 times). 152 avalanches are said to have been seen from the road.
- 58 persons in 30 different avalanches were mentioned to have been buried, some totally, others only partly. 67 persons were injured and 13 avalanche accidents ended for 25 people fatal.

7.1.2 Data Quality

7.1.2.1 Accuracy

Locational outliers in the avalanche point data set are shown in figure A.1. According to their coordinates 17 avalanches lie outside of Scotland, mainly in the sea. These must be typing errors and are not included in any further data analysis. 58 avalanches show sensible coordinates lying within Scotland but are located outside of the five SAIS areas.

The measurement imprecision or inaccuracy of variables is biggest within the attribute 'Region'. Originally, only 560 entries were classified correctly (17% of the total avalanches) and three were wrong (less than 1% of the total avalanches).

A comparison of the terrain attributes in the original data file to those from the DEM are shown in figure 7.10. The largest differences are within the attribute 'Gradient', as there are also by far the largest differences in the amount of available data (see table 7.6). Ranges at each number are high and the peak shifts from the interval (35,40] to (30,35]. Distributions in the altitude and aspect are similar, and the relationships are rather linear which speaks for a higher accuracy. The peak in the altitude shifts from 900 masl to 1100 masl in the original data set to 1000 masl to 1100 masl in the data of the DEM. A tendency for high frequencies at round numbers emerges in both attributes. Some values classified as northern to northeastern aspects in the original data set are rather classified as Northern to Northwestern aspects in the data of the DEM.

Gross errors due to obvious typing errors make up 0.4% of the total data set and are present in only three attributes (see table 7.3). Additionally, in some cases where exceptionally small or large numbers are present, it is not easy to decide if an entry is a typing error or not (see figure 7.4).

Table 7.3: Typing errors in the original data set of the reported avalanches.

Attribute	Typing Error	Date
Minimum Fracture Depth	-999	25.02.1997
Altitude	0.01	27.12.2001, 25.01.2002, 06.04.2004
	0.09	26.03.2004
Aspect	-1	1.2.2012
	0.02	07.02.2009
	0.03	24.01.2009, 29.12.2009
	0.05	07.03.2007
	0.45	03.01.2007
	445	06.01.2005
	525	08.03.1997
	980	11.02.1997, 16.02.2000
	1000	23.02.2002
	1200	27.02.2011

7.1.2.2 Resolution

Spatial resolution of the DEM is relatively low with 50 meters and influences the information that is extracted, especially the gradient (Zhang *et al.* 1999, Kienzle 2004, Thompson *et al.* 2001). This is taken into account when interpreting the data.

Variable resolution seems to be appropriate for most attributes. Fall lengths and widths are given in meters, fracture depth in centimetres. The attributes 'Type', 'Release' and 'Region' are on a scale which has been defined by the SAIS. 'Size' is given on a frequently used scale reaching from 1 to 5. In some attributes, it is not simple to decide without any doubt if specific numbers are correct or given in an other unit (see table 7.4). For example, 11 entries in the attribute 'Minimum Fracture Depth' lie at 0.1, 0.25, 0.3 or 0.5. Since [mm] would be unrealistically small, these entries are probably meant in [m]. Also within the attribute 'Maximum Fracture Depth' 18 entries are included which are smaller than 1 and probably given in [m] instead of [cm]. For small numbers between 1 and about 5, it is even more difficult to decide on the correct unit.

7.1.2.3 Consistency

Internal consistency is ensured for the majority of attributes as the units do not change. Only in the attributes 'Fracture Depth', 'Fall Length' and 'Width' few entries are included which are unrealistically small and probably given in [m] instead of [cm] or [km] instead of [m] (see table 7.4). However, the amount of errors remains below 1% in any case.

For missing data, different labels are being used: 'NULL', '-9999', Empty cells and in the comments 'No details'. These can clearly be recognised as indications of missing data, but with '0' it

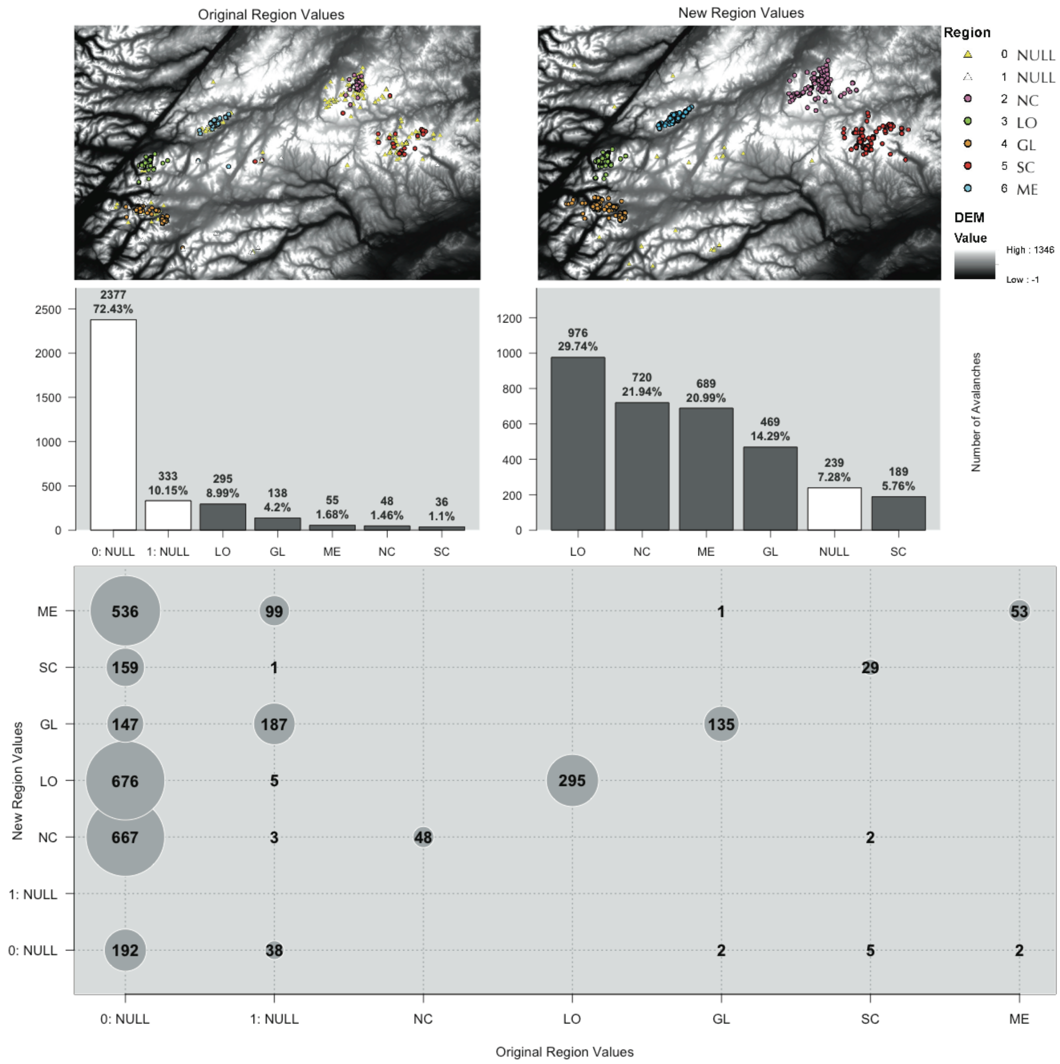


Figure 7.9: Comparison of original and new values of the attribute 'Region'.

is more difficult. In some cases, it probably represents missing data but has to be used as zero in others. Since no metadata is available for the data set, '0' is interpreted as missing value if zero is not sensible or the amount of '0' is not realistic (see table 7.7).

Consistency of related attributes is tested for minimum and maximum width and fracture depth (see table 7.5). The row 'MIN larger than MAX' shows the amount of avalanches in which consistency is not ensured, either because the numbers are mixed up or because typing errors are present. In both attributes the inconsistency is low, affecting only 0.2% of the total data.

7.1.2.4 Completeness

Spatial coverage of the DEM is sufficient as all five SAIS areas are fully covered (see figure A.1).

Variable values which are missing are the most important measure for completeness and amounts are summarised in table 7.7 for each attribute. 'Date' is the only attribute which is available for each avalanche. Only few missing values are present in the attributes 'Comment' and 'Altitude' as well as in the coordinates. More than 90% missing numbers occur for the attributes 'Size', 'Gradient' and 'Release Attempts'. In average about 54% of the data is missing in each attribute.

Variable values of the attribute 'Region' are newly assigned to each avalanche, so that the amount of available data is raised to about 93% (see figure 7.9).

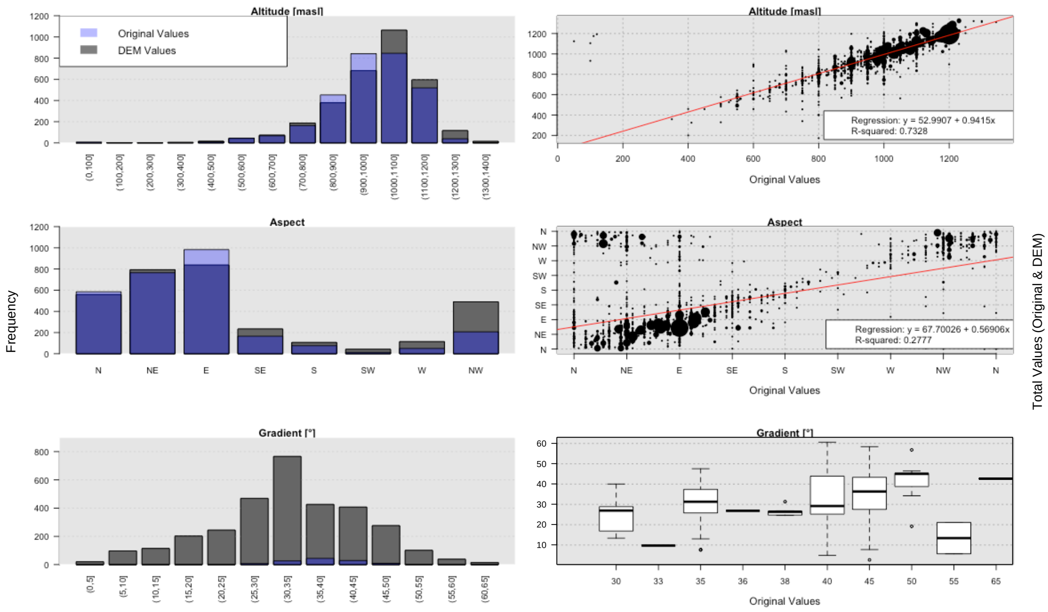


Figure 7.10: Distributions and differences of the original values and the total (original & DEM) values in the attributes 'Altitude', 'Aspect' and 'Gradient'.

Table 7.4: Attribute values in the reported avalanche data set which are possibly given in a wrong unit.

Attribute	Correct Unit	Unsure Values (Nr)	Total [abs: %]
Minimum Fracture Depth	[cm]	0.1, 0.25, 0.3, 0.5 (7) unrealistic 1 (2), 2, 4, 5 (13), 6 (1), 9 (1) small, but realistic	10: 0.3% 19: 0.58%
Maximum Fracture Depth	[cm]	0.03, 0.04, 0.1, 0.2 (3), 0.25, 0.3, 0.4, 0.5, 0.7 unrealistic 1 (13), 1.5 (2), 3 (3), 5 (4), 9 (2) small, but realistic	11: 0.33% 24: 0.73%
Minimum Width	[m]	0.02, 0.5 unrealistic	2: 0.06%
Maximum Width	[m]	0.5 unrealistic 1, 2, 3 (7), 4 (9) small, but realistic	1: 0.03% 18: 0.55%
Fall Length	[m]	0.04 unrealistic 2, 5 (6) , 6, 7 (2), 8 (3) small, but realistic	1: 0.03% 13: 0.40%

Table 7.5: Comparison of minimum and maximum values for the attributes 'Width' and 'Fracture Depth'.

	Width		Fracture Depth	
	Number of Values [abs]	Number of Values [%]	Number of Values [abs]	Number of Values [%]
Neither MIN Nor MAX	2512	76.5	1763	53.7
Only MIN	36	1.1	192	5.9
Only MAX	427	13.0	612	18.6
MIN Equal to MAX	80	2.4	294	9.0
MIN Smaller than MAX	217	6.6	407	12.4
MIN Larger than MAX	5	0.2	8	0.2

7.1.2.5 Model Quality

Spatial object representation of avalanches should most accurately be done using polygons, because avalanches are four dimensional, including spatial coordinates, volume and time. To make it even

Table 7.6: Amount of available information in the original data, in the data of the DEM and the total combined data for the attributes 'Altitude', 'Aspect' and 'Gradient'.

	Altitude		Aspect		Gradient	
	[abs]	[%]	[abs]	[%]	[abs]	[%]
Total Original Data Set	3002	91.47	2958	90.13	153	4.66
Total New (Original & DEM)	3062	93.30	3062	93.30	3062	93.30
Identical Values	70	2.13	3	0.09	0	0.00
DEM > Original	1229	37.45	1326	40.40	25	0.76
DEM < Original	1560	47.53	1363	41.53	82	2.50
NULL in Original, Value in DEM	129	3.93	301	9.17	2888	88.00
NULL in DEM, Value in Original	143	4.36	266	8.10	11	0.34
NULL in Both	151	4.60	23	0.70	276	8.41

Table 7.7: Amount of missing data in the original data set of the reported avalanches.

Attribute	'-9999'	'NULL'	Empty Cell	'0'	'0' as 0 or NULL?	Total [abs]	Total [%]
Date	-	-	-	-	-	0	0.00
Comment	-	-	123 (+7: "No details")	-	-	130	3.96
Easting/Northing	164	-	-	-	-	164	5.00
Longitude/Latitude	-	178	-	-	-	178	5.42
Altitude	54	7	-	219	NULL	280	8.53
Aspect	318	6	-	272	0	324	9.87
Fall Length	1070	67	-	471	NULL	1608	48.99
Maximum Width	1086	173	-	695	NULL	1954	59.54
Region	-	-	-	2377	NULL	2377	72.43
Minimum Width	1291	409	-	680	NULL	2380	72.52
Release	1293	396	-	766	NULL	2455	74.80
Avalanche Type	1693	40	-	778	NULL	2511	76.51
Maximum Fracture Depth	1399	248	-	903	NULL	2550	77.70
Hour	-	-	-	2812	2801: NULL, 11: 0?	2812	85.68
Minimum Fracture Depth	1475	445	-	1024	NULL	2944	89.70
Size	-	-	-	2981	NULL	2981	90.83
Gradient	-	325	-	2804	NULL	3129	95.34
Release Attempts	1640	640	-	1002	NULL	3282	100.00

more complex, avalanches are involved in a process and not stationary. Therefore the representation of an avalanche as a point is a strong simplification. Estimates about exact numbers such as distance from ridges are not possible, it cannot be distinguished between starting zone or flow path and magnitudes cannot be included in the spatial objects.

However, "if analysis can disregard internal differentiation and if the geographical scale of the analysis means that detailed estimates of spatial relationships such as distance or areal configuration are not needed, then the representation of an areal object [...] by a point is appropriate" (Haining 2003, 60). This is true for the application in this thesis. The scale of the current analysis concerns not only one avalanche but is rather at different areas or even Scotland, and the detailed extent of one single avalanche does not play a decisive role. Therefore a representation as points is sufficient. Additionally, some of the above mentioned properties are already included in attributes.

Attribute completeness seems to be ensured. In contrast, the attribute 'Release Attempts' could rather be deleted as 100% of the attribute values are missing.

7.1.2.6 Summary

To sum up, the avalanche data set shows characteristics that limit the data quality, especially considering the amount of missing data and uncertainties due to consistency issues. But considering the scope of this thesis, this is not a limiting factor as exactly the detection of such implications are part of the main aim. However, data quality issues are influencing the description of typical avalanche characteristics. The results therefore have to be considered critically and the data quality aspects have to be kept in mind.

7.2 Part B: Combination of Data Sets

7.2.1 Overview

The different data sets are combined and two data sets result. The first one includes the snow profile header information, the second one the depth information.

Header Information
To reach a maximum temporal coverage for each header attribute in every area, the following attributes are added from other data sources to the *internet data*:

from ProfileObs: Precipitation Code, Observed Avalanche Hazard, Forecast Avalanche Hazard, Observer, Grid, Altitude, Aspect, Incline, Location, Total Snow Depth, Ski Pen, Comments

from CorniceWx: Snow Index, No Settle, Insolation, Crystals, Wetness, Avalanche Category, Snow Temperature, Rain

from Binary files, header information: Precipitation Code, Observer, Observed Avalanche Hazard, Forecast Avalanche Hazard, Grid, Altitude, Aspect, Incline, Location, Total Snow Depth, Ski Pen, Comments

from SIESAWS: Summit Air Temperature, Summit Wind Direction, Summit Wind Speed

Figure 7.12 shows a graphical overview of this puzzle. Missing data is indicated with white lines.

Depth Information
The resulting depth information consists of the data of the *binary files* up until 2007, and between 2007 and March 2013 of *Book1* data. Each snow profile contains several rows which describe layers at different depths.

7.2.2 Data Quality

The data quality of the original and the resulting data sets described in the last section is analysed. If possible, characteristics which limit data quality are removed in the resulting data set to increase the data quality. Both steps are described below.

7.2.2.1 Accuracy

The measurement imprecision is analysed only by comparing related attributes, as no ground truth data is available. Figure 7.11 shows the relation between weather measurements at the snow profile locations and at automatic weather stations for the attributes 'Air Temperature', 'Wind Direction' and 'Wind Speed'.

Relationships in the air temperature and wind speed are close to linear and only little entries show unrealistically large differences between the two measurements. Summit measurements are, especially at very low temperatures, on average about 2 °C lower. Wind direction values show larger differences, but as they are highly variable depending on the specific terrain characteristics (e.g. Price 2000, Roy 1983), large differences do not have to be imprecision. In all three attributes, rounding influences in the snow profile data are visible.

Gross errors in the original data set are mainly typing errors which are summarised in table 7.8 and removed in the resulting data sets. Most typing errors are present in the attributes 'Crystals', 'Precipitation Code' and 'Wetness', but are with less than 1% of the total data only few.

Table 7.8: Obvious typing errors which are removed in the combined snow profile data set.

Attribute	Typing Error	Area	Amount [abs]	Amount Total [abs] / [%]	Dates
Altitude	25885	SC	1	1 / 0.01	18.02.11
Aspect	390	ME	1	1 / 0.01	15.03.11
Cloud	199	ME	1	1 / 0.01	25.02.97
Incline	955	NC	1	1 / 0.01	18.02.11
Summit Wind Speed	-8	LO	1	1 / 0.01	06.01.10
Wind Speed	-2	ME	1	1 / 0.01	23.12.10
Avalanche Cat	-2	ME	1	7 / 0.06	18.01.11
	88	GL	2		22.03.13, 23.03.13
	99	GL	1		19.02.11
	880	NC	1		08.12.11
	1011	NC	1		15.02.11
	4400	GL	1		31.01.11
Aval Code	77777	LO	1	10 / 0.09	25.01.02
	44000	GL	1		21.02.05
		SC	1		30.01.09
		ME	4		12.01.06, 18.02.11, 11.01.06, 30.01.09
	99000	SC	1		31.03.10
		ME	1		07.03.09
	10041	SC	1		13.01.11
Snow Index	1	NC	1	12 / 0.11	20.11.10
	3	ME	1		18.02.11
	5	NC	1		13.02.10
	18	NC	1		02.01.09
	20	NC	1		18.01.06
	24	SC	1		20.12.10
	26	LO	1		20.12.11
	90	LO	1		13.01.13
	170	LO	1		24.02.12
	172	LO	1		09.02.13
	210	NC	1		07.02.13
	282	GL	1		12.04.13
Insolation	1	ME	1	27 / 0.24	23.12.05
	5	NC	1		19.01.06
		LO	1		10.02.11
	13	GL	1		20.03.12
	15	NC	1		04.04.12
		GL	2		04.05.13, 05.05.13
		ME	1		03.04.12
	24	NC	5		16.04.91, 17.04.91, 18.04.91, 19.04.91, 16.04.95
		LO	8		16.04.91, 17.04.91, 18.04.91, 19.04.91, 20.04.91, 16.04.92, 17.04.92, 16.04.06
		GL	5		16.04.92, 17.04.92, 18.04.92, 19.04.92, 16.04.95
	208	LO	1		07.03.13
Crystals	1	NC	12	63 / 0.56	04.04.11, 11.12.10, 20.01.11, 14.02.12, 14.03.12, 15.03.12, 21.03.12, 09.01.13, 10.01.13, 31.01.13, 05.02.13, 03.05.13
		LO	36		04.01.11, 05.01.11, 10.01.11, 12.01.11, 13.01.11, 19.01.11, 20.01.11, 21.01.11, 25.01.11, 26.01.11, 27.01.11, 28.01.11, 01.02.11, 02.02.11, 03.02.11, 07.02.11, 08.02.11, 10.02.11, 11.02.11, 12.02.11, 13.02.11, 22.02.11, 24.02.11, 25.02.11, 17.03.11, 23.01.12, 26.01.12, 18.02.12, 25.02.12, 03.04.12, 25.01.13, 29.01.13, 31.01.13, 04.02.13, 05.02.13, 10.02.13
		SC	9		28.01.11, 29.01.11, 13.02.11, 14.02.11, 15.02.11, 16.02.11, 17.02.11, 22.02.11, 14.03.11
		ME	1		21.01.11
	2	NC	1		29.03.07
		SC	4		24.12.11, 26.12.11, 10.01.12
Precipitation Code	1	NC	10	85 / 0.76	02.04.00, 10.02.01, 12.03.01, 23.02.02, 25.02.02, 07.03.02, 25.01.02, 10.03.02, 03.02.02, 01.03.02
		LO	14		22.12.95, 17.01.98, 08.04.98, 24.12.99, 04.01.99, 01.01.00, 28.02.00, 05.03.00, 26.01.01, 25.01.02, 03.02.02, 18.02.02, 10.03.02, 11.03.02
		GL	20		14.04.92, 08.01.93, 29.12.93, 26.12.98, 12.01.99, 21.01.99, 23.01.99, 25.01.99, 31.12.00, 24.01.01, 06.02.01, 07.02.01,10.02.01, 03.03.01, 25.01.02,18.02.02, 20.02.02, 21.02.02, 23.02.02, 10.03.02
		SC	5		02.04.00, 04.02.01, 23.02.01, 25.01.02, 10.03.02
		ME	34		07.03.98, 12.04.98, 23.12.98, 24.12.98, 12.01.99, 15.01.99, 17.01.99, 07.02.99, 27.02.99, 14.03.99, 09.02.00, 10.02.00, 12.02.00, 13.02.00, 28.02.00, 24.01.01, 26.01.01, 06.02.01, 25.01.02, 05.02.02, 20.02.02, 22.02.02, 23.02.02, 10.03.02, 11.03.02, 02.02.03, 08.02.04, 08.01.05, 18.01.05, 09.02.05, 10.02.05, 13.02.05, 12.03.06, 23.03.06
	3	LO	1		30.12.01
	5	GL	1		26.12.95
Wetness	3	NC	5	85 / 0.76	24.01.01, 14.12.04, 29.12.04, 22.03.05, 24.03.05
		LO	34		13.01.06, 14.01.06, 19.01.06,12.01.07, 13.01.07, 15.02.07, 04.03.07, 10.03.07, 11.03.07, 16.03.07, 17.03.07, 23.03.07, 24.03.07, 31.03.07, 01.04.07, 02.04.07, 03.04.07, 04.04.07, 05.04.07, 06.04.07, 07.04.07, 31.12.10, 13.01.11, 24.01.11, 25.01.11, 31.01.11, 02.02.11, 25.02.11, 16.03.11, 22.03.11, 31.03.11, 22.12.11, 26.12.11, 28.12.11
		ME	7		31.12.11, 04.01.12, 10.01.12, 20.02.12, 23.12.12, 03.03.13, 03.05.13
	4	GL	15		02.01.03, 11.02.06, 19.03.06, 26.12.06, 13.01.07, 18.01.07, 04.03.07, 05.03.07, 18.03.07, 17.01.08, 31.01.08, 20.02.08, 01.03.08, 12.03.08, 26.03.08
		ME	6		23.02.03, 14.01.11, 16.01.11, 28.12.12, 31.12.12, 07.01.13
	5	ME	1		16.03.05
	8	GL	17		07.02.03, 15.02.06, 09.03.06, 23.12.06, 23.02.07, 06.03.07, 27.12.07, 29.01.08, 30.01.08, 03.02.08, 22.02.08, 26.02.08, 05.03.08, 07.03.08, 08.03.08, 01.04.08, 02.04.08

7.2.2.2 Resolution

Spatial resolution in the snow profile data is such that for any of the five SAIS areas one snow profile is available. This is not a lot in comparison to all locations at which avalanches release and especially to the high variability of conditions. However, the locations of snow profiles are chosen carefully by the SAIS. Each snow profile should represent the most hazardous conditions of the area and describe a worst case scenario (Diggins 2010). In general, the snow profile information therefore should be applicable for the whole area, but this has to be done carefully as the information is indeed only measured at one single location on one certain slope.

The summit attributes are measured at automatic weather stations which are installed at fixed locations. They are exposed, at high altitudes, and tend also to be rather extreme.

Temporal resolution is on a daily scale, as in each area one snow profile is taken per day. The data measured at snow profile locations describe the situation at the time of the recording and summit

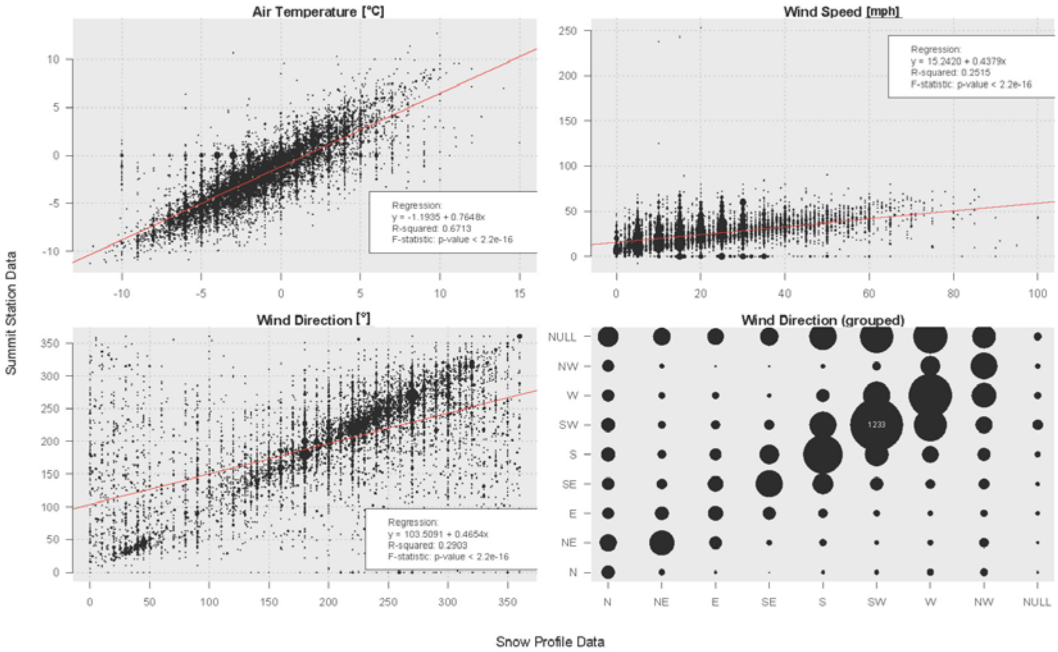


Figure 7.11: Relations of the measurements taken at the snow profile locations and at the summit stations for the attributes 'Air Temperature', 'Wind Direction' and 'Wind Speed'.

measures are averages over the whole day. The snow profile information therefore can be used as typical for the whole day, only depending on the temporal variability of the specific attribute. In case of high short-term variability, the measurements can be misleading.

Variable resolution generally seems to be well chosen and only few important issues have to be considered. For air and snow temperatures, the resolution differs within the attribute as values are partly given in [1/10 °C] and in [°C]. For general applications, [°C] should be sufficient and measurements in the field at an accuracy of [1/10 °C] need very careful work and precise instruments.

The attributes 'Precipitation Code' and 'Snow Index' are relative measurements of the amount of new snow including a 6-step scale reaching from *'no snow'* to *'very heavy snowfall'*. The actual number is dependent only on the perception of the observer and therefore is subjective.

7.2.2.3 Consistency

Internal consistency is ensured as the units do not change for the majority of attributes, except for the air and snow temperature[1]. The attribute 'Air Temperature' is in the *internet files* for the Northern Cairngorms, Glencoe and Creag Meagaidh before the year 2003 partly given in [1/10 °C] and partly in [°C] (see figure 7.12, first row). No general pattern in the distribution of the two different units is found. Neither dividing all values by ten (see figure 7.12, second row) nor dividing only numbers above 10 °C and below -10 °C by ten (see figure 7.12, third row) led to a satisfying result.

Therefore the *internet file* data is adjusted manually to reach maximum consistency. *ProfileObs* data is used as reference. As the air temperature in the other data sets is provided in [°C], values given in [1/10 °C] are adjusted considering the following rules:

- NC: numbers in intervals [-120, -20] and [16, 116] are divided by 10.
- GL: numbers in intervals [-110, -11] and [11, 126] are divided by 10.
- ME: numbers in intervals [-90, -11] and [11, 140] are divided by 10.
- Remaining numbers are divided manually, if the divided number is closer to the *ProfileObs* value than the original one.
- If no information is available in the *ProfileObs* data, the values are not changed.
- Numbers with a single digit are not divided.
- The entry '25085' in the SC data is deleted as it is an obvious typing error.

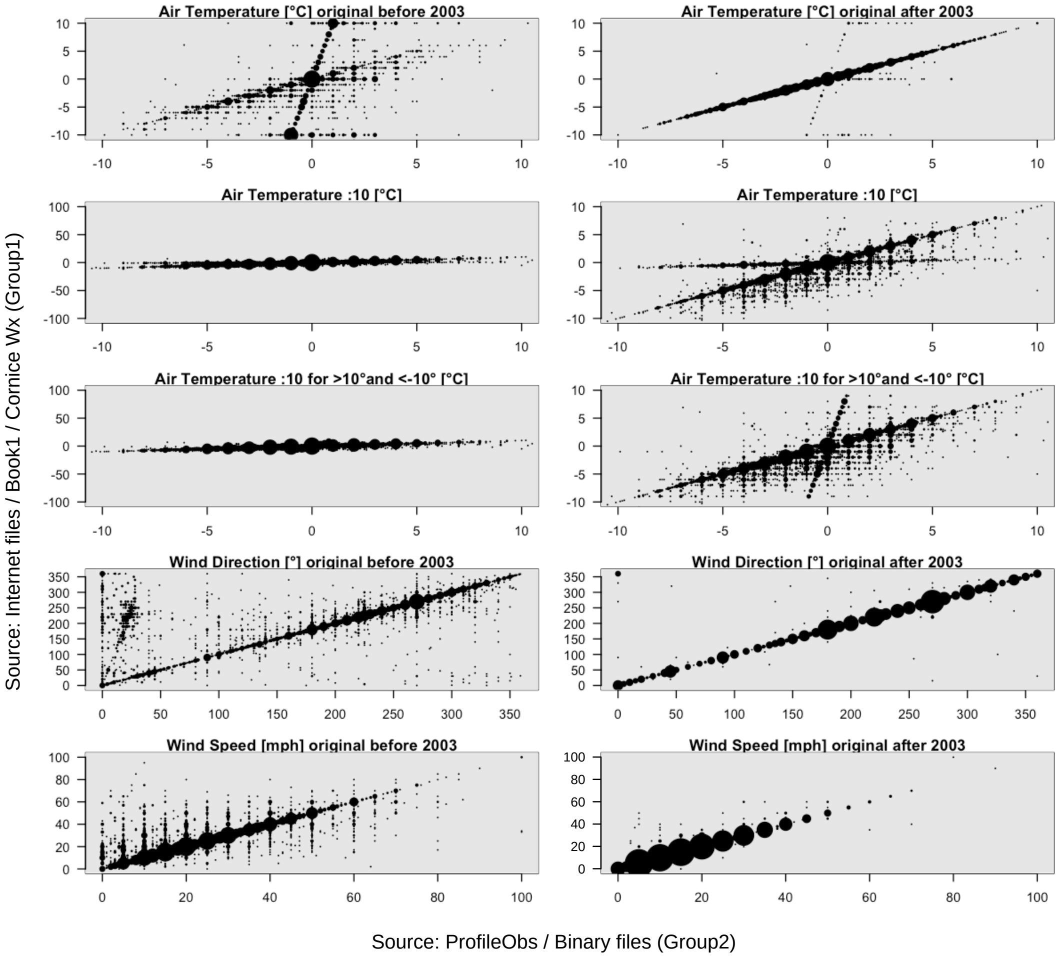


Figure 7.12: Comparison of the attribute values 'Air Temperature', 'Wind Direction' and 'Wind Speed' from different data sources.

Missing data is indicated differently with '-9999', 'NULL', 'N/A' or empty cells. Sometimes also '0' has to be counted as missing data, depending on its sensibility (see table 7.10).

Further, the wind speed and wind direction entries seem to be mixed up in some cases. Wind speeds above 100 mph are very seldom in Scotland and winds come mostly from W and S (see figure 8.2, Buchan 1890, MetOffice 2013c). Therefore the values in the resulting data set, where wind speeds are bigger than 100 and wind directions below 100, are exchanged. With unrealistic high wind speeds but plausible wind direction, only the wind speed values are deleted as they are probably rather typing errors (see table 7.9 for amount). Similar issues were already found by Purves *et al.* (1998b).

Consistency when linking data sets can be assessed for a nearly infinite number of data pairs, because many attributes are included in various data sets. Not every step is presented here as it would go too far, but the following statements summarise the analysis:

- **Group1**: the values of the data sources ***Book1***, ***Internet files*** and ***CorniceWx*** are identical.
- **Group2**: the values of the data sources ***ProfileObs*** and ***Binary files*** are identical.
- The values of **Group1** and **Group2** are identical except for the attributes 'Air Temperature', 'Wind Direction' and 'Wind Speed' for the years before 2003 (see figure 7.12).

Differences between the two groups are in all three attributes larger before the year 2003. For the attributes 'Wind Direction' and 'Wind Speed', a considerable amount of the differences seems

[1]Snow temperature is newly assigned to each snow profile in any case and is not assessed further.

Table 7.9: Number of values in the attributes 'Wind Speed' and 'Wind Direction' that are mixed up.

Area	Wind Speed Number Switched	Wind Speed Number Deleted	Summit Wind Speed Number Switched	Summit Wind Speed Number Deleted
Northern Cairngorms	0	2	19	1
Lochaber	7	0	14	2
Glencoe	1	1	5	0
Southern Cairngorms	6	0	16	0
Creag Meagaidh	4	0	5	0

to be due to rounding issues in the *ProfileObs* or *binary file* data. This is little important as these values are only seldom used in the resulting data set. Also small wind directions below 50 seem to be rather errors in the *ProfileObs* data as those values are seldom in Scotland (see figure 8.2, Buchan 1890, MetOffice 2013c). They are probably typing errors with two instead of three digits.

Special consideration is given to the summit data as these attributes are only available since 2007 in the *internet files* and play an important role in part C of this thesis. The comparison for the years 2007 to 2013 for the Northern Cairngorms between the *internet files*, *SIESAWS* data and *Cairn Gorm* data is used to assess the consistency and to decide on which data source is used for the enhancing of the resulting data set. Best matching is found between the *internet files* and the *SIESAWS* data (see figure 7.13). As for the air temperature only minimum and maximum values are stored in the

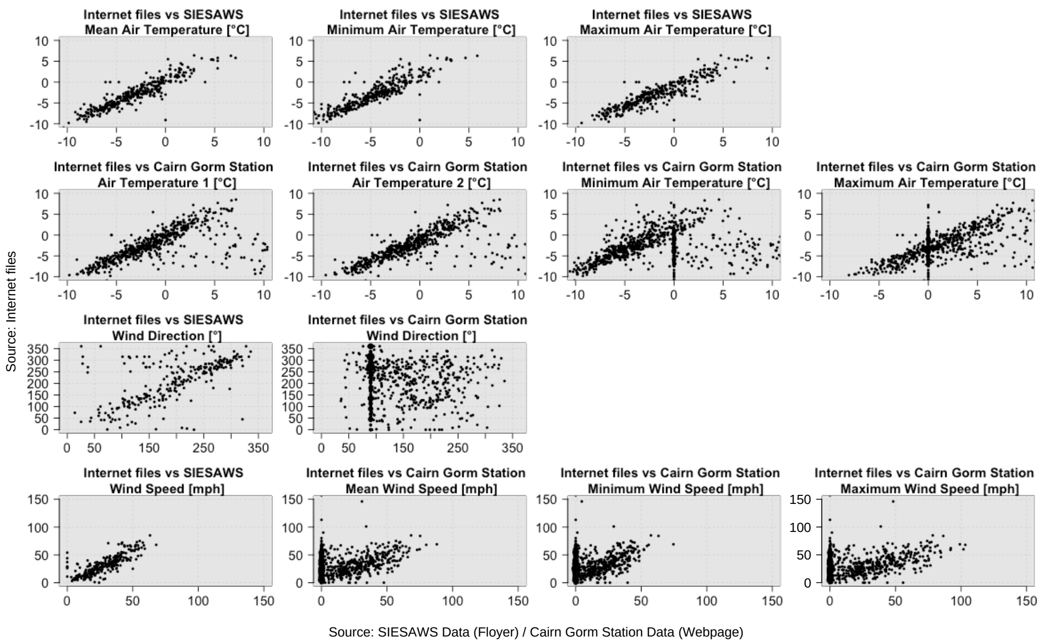


Figure 7.13: Comparison of the attributes 'Summit Air Temperature', 'Summit Wind Direction' and 'Summit Wind Speed' in different data sources for the Northern Cairngorms.

SIESAWS data, the mean of these two measures is calculated and used in the end. Therefore the *SIESAWS* data of Cairn Gorm is used for the resulting data set for the years 1992 to 2007. This is a good sign for the other areas, where only *SIESAWS* data is available. For Lochaber and Glencoe, data of Aonach Mor for the years 1992 to 2007 and for the Southern Cairngorms, data of Cairnwell for the years 1996 to 2007 is used.

Creag Meagaidh is located between the other SAIS areas and the observers use partly data of Cairn Gorm (1245 masl) and from Aonach Mor (1130 masl). The first two columns in figure 7.14 show the relation between the data included in the *internet files* and the *SIESAWS* data after 2007.

Since wind direction and wind speed are factors for which both data sources are available only on few days, a verification is limited possible. The last column shows a comparison between the enhanced data using either the information from Aonach Mor or from Cairn Gorm. Summit air temperatures are related nicely linear, but summit wind direction and wind speed show larger differences. Therefore the air temperature seem to vary less with distance than wind speed or wind direction. Since it is difficult to say which data is more accurate, both options are included in the resulting data set. In part C of this thesis, data of Aonach Mor is used, as this station lies closer to the Creag Meagaidh area, but data of Cairn Gorm are considered in interpretations.

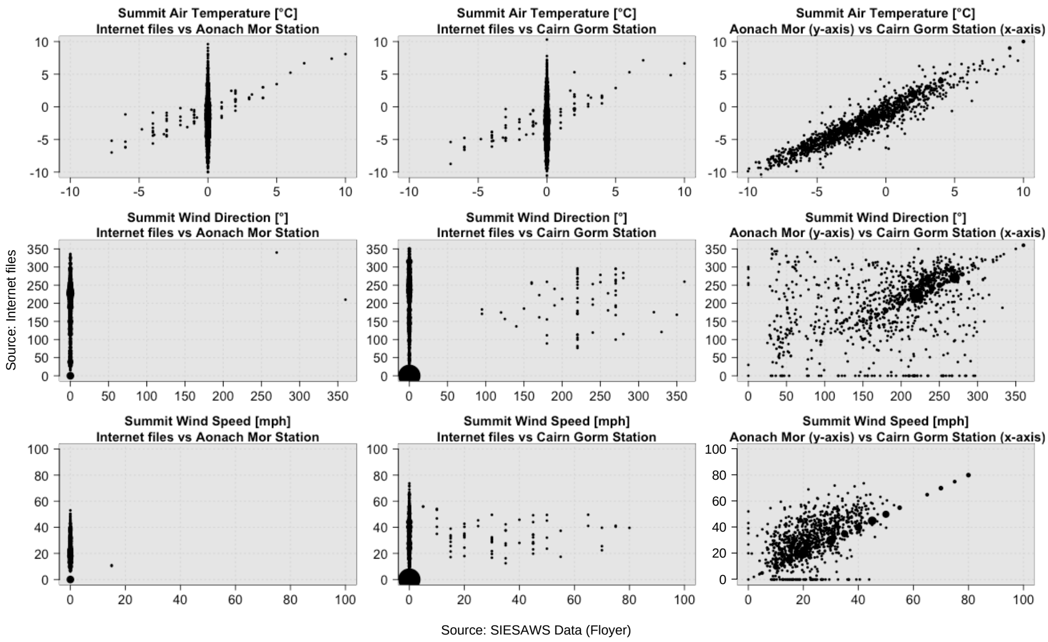


Figure 7.14: Comparison of the attributes 'Summit Air Temperature', 'Summit Wind Direction' and 'Summit Wind Speed' of stations at Aonach Mor and at Cairn Gorm for Creag Meagaidh.

7.2.2.4 Completeness

Spatial coverage is closely related to the spatial resolution and is generally ensured as snow profiles are taken in each of the five areas. As they are located at most hazardous conditions, they should be representative for the whole area (Diggins 2010).

Variable values are missing in the majority of attributes (see table 7.10). The largest amount of missing data show the attributes 'Ski Pen', 'Maximum Temperature Gradient', and 'Maximum Hardness Difference'. With the 'Foot Pen' is a similar attribute to the ski pen included where the amount of missing information is only 7%. The maximum temperature gradient and maximum hardness difference are included in the depth information as well and are newly calculated for part C of this thesis in any case. In average 24.4% of the data is missing per attribute.

Temporal coverage is used as basis for the combination of different data sources into the resulting data sets and is a crucial aspect of the data quality. Theoretically, the data sets should cover the entire seasons, usually lasting from November or the beginning of December to April or the beginning of May for every season in which the SAIS has been operating in the corresponding areas. The maximum coverage would be therefore from 1990 to 2013 for the Northern Cairngorms, Lochaber and Glencoe and from 1996 to 2013 for the Southern Cairngorms and Creag Meagaidh. As different data sets are available, the temporal coverage has to be examined for each data set, considering every attribute and area separately (see tables 7.11 for a detailed overview).

The total temporal coverage present in the resulting data file is maximised for each attribute in every area. Figure 7.12 shows an overview of the resulting data set where white spaces indicate still missing data. For the first two seasons in the very early beginning of the SAIS operation and for the years between 2003 and 2007, not all attributes are continuously available. Especially for the Cairngorms, a hole of four seasons in the majority of attributes exists. Creag Meagaidh is the only area in which the data is complete since operation start in 1996.

Considering part C of this thesis, a complete temporal coverage would particularly mean that for every day on which at least one avalanche is reported, snow profile information is available, preferably also for the days preceding the actual avalanche release. This can be ensured for all avalanche days but 23, on which in total 32 avalanches are reported (see appendix A.6.3 for exact dates).

Table 7.10: Amount of missing data in the combined snow profile data set. (Total number of snow profiles per area: NC: 2685, LO: 2646, GL: 2507, SC: 1373, ME: 1912, in total 11123 snow profiles.

Attribute	0	NC				LO				GL				SC				ME				TOTAL	
	as 0 or NULL?	NULL	0	total [abs]	total [%]	NULL	0	total [abs]	total [%]	NULL	0	total [abs]	total [%]	NULL	0	total [abs]	total [%]	NULL	0	total [abs]	total [%]	total [abs]	total [%]
Area id	-	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0		0	0.0	0	**0.0**
Status	-	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	**0.0**
Date	-	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	0	0	0.0	0	**0.0**
Drift	0	6	1222	6	0.2	2	1458	2	0.1	2	1210	2	0.1	12	846	12	0.9	0	997	0	0.0	22	**0.2**
Air Temperature	0	10	93	10	0.4	16	97	16	0.6	13	89	13	0.5	16	59	16	1.2	15	110	15	0.8	70	**0.6**
Cloud	0	7	115	7	0.3	11	156	11	0.4	12	74	12	0.5	28	34	28	2.0	15	59	15	0.8	73	**0.7**
Wind Speed	0	21	20	21	0.8	11	70	11	0.4	14	31	14	0.6	65	36	65	4.7	22	22	22	1.2	133	**1.2**
Snow Index	0	13	1059	13	0.5	9	1441	9	0.3	17	1134	17	0.7	69	767	69	5.0	86	730	86	4.5	194	**1.7**
Wind Direction	0	27	12	27	1.0	39	60	39	1.5	26	23	26	1.0	77	19	77	5.6	38	18	38	2.0	207	**1.9**
No Settle	0	66	31	66	2.5	7	47	7	0.3	4	27	4	0.2	339	14	339	24.7	340	20	340	17.8	756	**6.8**
Foot Pen	NULL	16	142	158	5.9	10	301	311	11.8	27	116	143	5.7	51	79	130	9.5	19	32	51	2.7	793	**7.1**
Snow Temperature	0	72	717	72	2.7	17	989	17	0.6	18	1005	21	0.8	355	303	355	25.9	396	782	396	20.7	861	**7.7**
Insolation	0	66	0	66	2.5	14	331	14	0.5	15	4	15	0.6	352	1	352	25.6	526	0	526	27.5	973	**8.7**
Wetness	0	72	1	72	2.7	22	2303	22	0.8	150	21	150	6.0	363	54	363	26.4	377	29	377	19.7	984	**8.8**
Crystals	0	75	1368	75	2.8	25	2493	25	0.9	220	2228	220	8.8	377	810	377	27.5	603	929	603	31.5	1300	**11.7**
Summit Air Temperature	NULL	290	0	290	10.8	488	40	528	20.0	449	43	492	19.6	47	0	47	3.4	598	6	604	31.6	1961	**17.6**
Rain	0	1934	648	1934	72.0	2	2083	2	0.1	2	1919	2	0.1	7	1243	7	0.5	16	1474	16	0.8	1961	**17.6**
Summit Wind Direction	0	479	0	479	17.8	505	7	505	19.1	552	10	552	22.0	159	16	159	11.6	597	8	597	31.2	2292	**20.6**
Obs	NULL	1072	0	1072	39.9	538	12	550	20.8	666	3	669	26.7	14	0	14	1.0	9	2	11	0.6	2316	**20.8**
Location	NULL	1072	12	1084	40.4	534	0	534	20.2	658	10	668	26.6	30	11	41	3.0	8	0	8	0.4	2335	**21.0**
Summit Wind Speed	NULL	416	68	484	18.0	508	7	515	19.5	551	10	561	22.4	164	16	180	13.1	593	8	601	31.4	2341	**21.0**
Altitude	NULL	1097	0	1097	40.9	584	3	587	22.2	660	3	663	26.4	19	0	19	1.4	14	0	14	0.7	2380	**21.4**
Forecast Avalanche id	NULL	885	0	885	33.0	669	8	677	25.6	740	0	740	29.5	23	0	23	1.7	91	0	91	4.8	2416	**21.7**
Grid	NULL	1083	0	1083	40.3	611	5	616	23.3	667	3	670	26.7	29	0	29	2.1	31	0	31	1.6	2429	**21.8**
Incline	NULL	1073	0	1073	40.0	635	11	646	24.4	673	5	678	27.0	84	5	89	6.5	33	0	33	1.7	2519	**22.6**
Observed Avalanche id	NULL	1078	0	1078	40.1	551	6	557	21.1	826	0	826	32.9	25	0	25	1.8	40	1	41	2.1	2527	**22.7**
Aspect	NULL	1075	34	1109	41.3	646	49	695	26.3	675	35	710	28.3	39	21	60	4.4	27	14	41	2.1	2615	**23.5**
Total Snow Depth	NULL	1089	0	1089	40.6	1187	231	1418	53.6	839	4	843	33.6	106	2	108	7.9	34	3	37	1.9	3495	**31.4**
Comments	NULL	1672	0	1672	62.3	1218	60	1278	48.3	1421	0	1421	56.7	491	0	491	35.8	455	3	458	24.0	5320	**47.8**
Time	NULL	1	1940	1941	72.3	0	16	16	0.6	0	1811	1811	72.2	0	685	685	49.9	0	1211	1211	63.3	5664	**50.9**
Avalanche Code	NULL	29	1711	1740	64.8	23	1637	1660	62.7	89	1557	1646	65.7	59	930	989	72.0	208	895	1103	57.7	7138	**64.2**
Precipitation Code id	NULL	1334	667	2001	74.5	597	1200	1797	67.9	900	850	1750	69.8	87	750	837	61.0	301	684	985	51.5	7370	**66.3**
AV Cat	NULL	2665	10	2675	99.6	1363	8	1371	51.8	1251	109	1360	54.2	836	4	840	61.2	1213	0	1213	63.4	7459	**67.1**
Max Hardness Difference	NULL	1939	11	1950	72.6	10	1920	1930	72.9	13	1829	1842	73.5	697	6	703	51.2	14	1215	1229	64.3	7654	**68.8**
Max Temperature Gradient	NULL	1130	336	1466	54.6	11	2364	2375	89.8	19	2255	2274	90.7	698	418	1116	81.3	1912	0	1912	100.0	9143	**82.2**
Ski Pen	NULL	2164	279	2443	91.0	1916	464	2380	89.9	2041	105	2146	85.6	1056	97	1153	84.0	649	1142	1791	93.7	9913	**89.1**

7.2.2.5 Model Quality

Attribute completeness is generally ensured as most important information about the location of the snow profiles, the weather and snow cover is included. It could be valuable to introduce two further attributes 'Rutschblock Score' and 'Shear strength'. This information is currently integrated in the comments and is tedious and time-consuming to extract.

Additionally, in several studies (e.g. Birkeland *et al.* 2001, Jomelli *et al.* 2007, McCollister *et al.* 2003 and Stoffel *et al.* 1998) as well as in the Observation Guidelines of Greene, E. (2010) (see sections 2.2.3 and 2.3), the snow water equivalent is considered as an important variable when analysing snow profile data. Such an attribute is not included in the current data and it might be useful to measure it additionally, to enable comparison to other studies.

7.2.2.6 Summary

To sum up, the snow profile data sets show various characteristics that limit the data quality. No data set exists, which can be seen as main source including the vast majority of information. Missing data, typing errors as well as inconsistencies between and within attributes affect the data most. But a considerable part of these limitations could be removed in the new combined data set and the data quality is raised in many aspects. Still, some limitations concerning mainly missing data exist and in some cases, it is difficult to assess the quality of attributes as no ground truth or other comparable measures are available.

7.3 Part C: Pattern Analysis

As no PCA is used for the definition of the input variables for the CA, several other considerations are used as basis.

7.3.1 Defining Variables

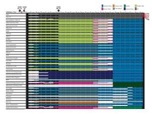
The following considerations lead to the choice of the input variables: a) The aim is to establish patterns describing situations that can vary. Therefore terrain attributes are not included. b) From the snow cover information, the general stability is the most important factor and a Weakness Index is introduced. c) Many different weather factors have an influence and they are included in the snow profile data. It is looked at each attribute to decide on its relevance:

- ***'Precipitation Code ID'*** and ***'Snow Index'*** include the same information, were in Part B combined into one single attribute, describe the amount of new snow and are very important factors.
- ***'Avalanche Hazards'*** are a consequence of the avalanche activity and do not influence the avalanche activity themselves.
- ***'Observer'*** and ***'Grid'*** are not relevant.
- ***'Terrain attributes'*** and ***'Location'*** only describe the specific location of the profile and cannot directly be related to high avalanche activity.
- ***'(Summit) Air Temperature'*** and ***'(Summit) Wind Speed'*** are very important factors.
- ***'(Summit) Wind Direction'*** only determines the aspects that are influenced by drifting but has no direct influence on the intensity of avalanche activity.
- ***'Air Temperature'***, ***'Wind Speed'*** and ***'Wind Direction'*** measured at the snow profile locations describe the situation at that specific slope at the time of recording. ***'Summit Air Temperature'***, ***'Summit Wind Speed'*** and ***'Summit Wind Direction'*** are measured at exposed locations and are averages over the whole day. They show rather extreme conditions, are more representative for a larger area as well as for the whole day and are therefore more appropriate.
- ***'Rain'*** and ***'Drift'*** are boolean attributes and are difficult to include into the CA.
- ***'Cloud'*** does not have such a direct influence on avalanche activity as, for example, air temperature or wind and ***'Insolation'*** includes a high amount of missing values.
- ***'No Settle'*** describes the amount of snow that has fallen since the beginning of the winter, has no direct influence on the specific conditions at a single day and is therefore little relevant for the current purpose.
- ***'Foot Pen'*** and ***'Ski Pen'*** include many missing values and describe mainly the amount of snow that is available for an avalanche. The stability is better described by other attributes which are included in the Weakness Index.

Table 7.11: Summary of the temporal coverage of each attribute in different data sources for each SAIS area.

[illegible]

Table 7.12: Overview of the combination of different data sources per attribute and SAIS area.



- ***'Total Snow Depth'*** is very dependent on the specific location of the snow profile and varies heavily concerning different altitudes, aspects and gradients. It can not be related to the location of the avalanche release easily and is therefore not relevant.
- ***'Avalanche Code'*** and ***'Avalanche Category'*** rather describe the avalanches themselves than the conditions at the time of release and are therefore not relevant.
- ***'Wetness'***, ***'Crystals'***, ***'Maximum Temperature Gradient'***, ***'Maximum Hardness Difference'*** and ***'Snow Temperature'*** are important factors that describe the properties of the snow cover.
- ***'Comments'*** mainly include additional information about properties of the snow cover and can therefore be relevant.

In summary, important attributes available in the data are 'Snow Index', 'Summit Air Temperature', 'Summit Wind Speed', the attributes describing the snow cover and the comments.

Similar studies mostly have used variables such as new snow, minimum and maximum air temperatures, snow water equivalent, total snow depth, wind speed and wind direction (e.g. Birkeland *et al.* 2001, Jomelli *et al.* 2007, McCollister *et al.* 2003). Especially for new snow, air temperatures and snow water equivalent often sums over several days preceding the avalanche day are calculated (e.g. Birkeland *et al.* 2001, Harvey *et al.* 2002).

Important conditions leading to high avalanche activity which are already known include thaws, storms and cold periods (Diggins 2010, Diggins 2011b, Diggins 2012, Ward 1980, Ward 1984a). These considerations lead to the definition of the following input variables for the CA:

- Summit air temperature at the avalanche day
- Sum of summit air temperature over the last three days preceding the avalanche day
- Sum of summit air temperature over the last seven days preceding the avalanche day
- Difference of summit air temperature between the avalanche day and the preceding day
- Number of days with negative air temperatures[2]
- Summit wind speed at the avalanche day
- Sum of summit wind speeds over the last three days preceding the avalanche day
- Sum of summit wind speeds over the last seven days preceding the avalanche day
- Snow index at avalanche day
- Sum of snow index over the last three days preceding the avalanche day
- Sum of snow index over the last seven days preceding the avalanche day
- Weakness Index including properties of the snow cover

Most of these variables have to be calculated from the existing attributes:

- For the attributes 'Summit Air Temperature', 'Summit Wind Speed' and 'Snow Index', the sum over the last three days and over the last seven days is calculated for every day, in case that they are directly following each other. These new variables take the history over the last few days and over the last week into account. Periods of strong snowfall, cold or warm periods and windy days can be considered.
- The difference for every day to the day before is calculated from the 'Summit Air Temperature'. Short time temperature changes, especially fast warming, can be accounted for.
- To include snow cover properties, a 'Weakness Index' is introduced and calculated from the snow profile depth information. Referring to the Nietentest (Schweizer 2007) and other proposed methods to measure snow cover stability in section 2.2.2.2, the sum of the variables shown in table 7.13 is used as Index. If the property is present in the snow profile, one point is added to the 'Weakness Index'. The higher the value of the 'Weakness Index', the weaker the snow cover tends to be.
  Crystal sizes are rounded because information such as '2-3' is often included in the data.
  All crystal forms except rounded grains and melt-freeze crust are seen as potential weak crystals. This decision is based on literature (see section 2.2.2 and table 2.4). Crystal forms are

[2]This variable was included in the first runs of the CA, but the influence on the resulting clusters is very high. Exceptionally long cold periods were put into one cluster which included only days of two or three consecutive weeks of a winter. Therefore this variable is excluded from the definitive CA.

Table 7.13: Variables and their thresholds included in the Weakness Index.

Variable	Threshold	Points
Maximum Crystal Size	>= 1 mm	1
Minimum Hardness	F / 4F	1
Number of Weak Crystal Forms	> 0	1
Maximum Crystal Size Difference	>= 3/4 mm	1
Maximum Hardness Difference	>= 2	1
Minimum Depth of Weak Crystals below Surface	< 1m	1
Maximum Temperature Gradient	> 1°/10 cm	1
Maximum Wetness Difference	>= 2	1
Minimum Number Resulting from Shear Tests	1, 2, 3	2
	4	1
Minimum Number Resulting from Rutschblock Tests	1, 2, 3	2
	4, 5	1
Weakness Index		**Sum**

very diverse, many transitions exist and they can act differently in relation with other factors. Some forms, including faceted crystals or depth hoar, are clear signs for instability, but other forms such as wet grains or ice masses do not have, in any case, a destabilising effect. Since weak crystals are only one factor out of ten that contribute to the 'Weakness Index', all crystal forms, mentioned by at least one author to be unstable, are considered.
Results from Shear tests and Rutschblock tests (see section 2.2.2.2) are included in the comments of the snow profiles. With keyword search in Excel the information is extracted and the result of the tests is assigned to the snow profiles. With the Rutschblock test, the numbers are directly given in the comments, the results of Shear tests are described in words including a variety of abbreviations. Using alphabetical sorting in Excel, they were manually transformed into numbers the following: 'very easy' = 1, 'easy' = 2, 'easy/moderate' = 3, 'moderate' = 4, 'moderate/hard' = 5, 'hard' = 6.

- For each day the number of days are counted that show temperatures either above or below 0°, depending on the air temperature of the current day. This enables the detection of long cold or long warm periods throughout the winter. The new variable 'Number of days with negative air temperature' results.

7.3.2 Cluster Analysis

The CA aims to group the avalanche days so that each cluster represents one typical situation in which a high avalanche activity is expected. Since the CA is a structure detecting method, no preliminary knowledge is necessary in the first place (Backhaus *et al.* 2008), but decisions on methodologies and parameters have to be made. These are described in the next sections.

7.3.2.1 Input Data

The input variables are stored in a data matrix as columns, the days as rows. Additionally, the date, area id and the number of avalanches are included which released at each day in total and per area, to allow clear selection, interpretation and referencing. As the CA is aimed to result in patterns for high avalanche activity, only days on which avalanches released are included.

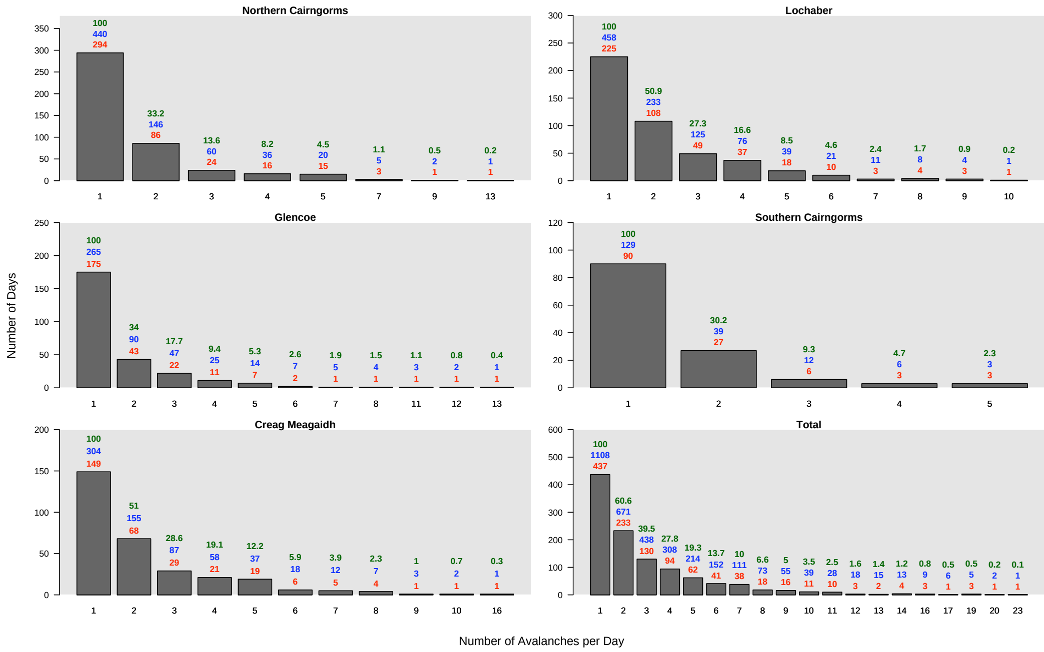


Figure 7.15: Number of days with different numbers of avalanches per day for all areas and in total. (red: absolute numbers, blue: reversed cumulative sum)

Definition of the Avalanche Day

The threshold used in the definition of an avalanche day as *a day on which a certain minimum number of avalanches occurred* is a crucial factor that determines the amount of input data. Figure 7.15 shows distributions of number of days against different numbers of avalanches per day for all five SAIS areas separately and in total. The blue numbers can be seen as the absolute amount of days and the green numbers as the percentage of days which are included in the input data when the threshold is set at the corresponding number of avalanches per day.

The decisions on the thresholds are based on several factors. In previous research, different approaches are found. Harvey *et al.* (2002) and Signorell (2001) used days with at least four avalanches as avalanche days, Birkeland *et al.* (2001) considered the upper 10% as avalanche days and Casteller *et al.* (2011) defined an Event Index and used the years which exceeded the long-term mean value. Further, running a CA with a certain number of resulting clusters makes only sense when a considerable amount of data is available and not only one or two cases are included in each of the resulting clusters. These considerations lead to the decision that the threshold for an avalanche day is chosen for each of the five SAIS areas separately, because the amount of available avalanche days varies considerably, especially for the Southern Cairngorms. For each area, several scenarios using different thresholds are calculated to find optimal solutions (see appendix A.9 for detailed plots).

Missing Data

As mentioned in section 6.2, different possibilities to deal with missing data exist. Omitting days on which at least one attribute is missing would be the simplest way, but as can be seen in table 7.14, the number of input data would at least be halved and a great amount of the available information not considered in the analysis[3]. Further possibilities include substitutions of missing values either with measures such as means or by estimating the values using models, for example a regression (Haining 2003).

As the aim of this thesis is to look at the available data, neither deleting whole days nor enhancing the data and introducing even more uncertainties into the analysis are seen as reasonable options. The remaining solution is to choose a clustering algorithm that can deal with missing values by

[3]Differences in the values between table 7.14 and figure 7.15 are due to days on which avalanches occurred but no snow profile data is available. See appendix A.6.3.

just ignoring them and using every value which is available for the analysis. However, to ensure a smooth run of the clustering, days in which only one to four values are available are not included in the analysis (see table 7.15 for exact amount).

Table 7.14: Amount of available data for different thresholds for an avalanche day, with or without omitting days with missing values.

Area	Snow Profiles	Input Variables		>=1 Avalanches / Day		>=2 Avalanches / Day		>=3 Avalanches / Day		>=4 Avalanches / Day		>=5 Avalanches / Day	
	Total	Total (>=0)	Omit	Total	Omit	Total	Omit	Total	Omit	Total	Omit	Total	Omit
NC	2685	1080	391	427	155	144	57	58	23	35	15	20	10
LO	2646	1066	697	453	277	233	142	125	73	76	38	39	16
GL	2507	1037	544	263	127	89	42	47	23	25	12	14	6
SC	1373	608	414	124	83	38	27	11	9	6	5	3	3
ME	1912	798	488	303	199	155	97	87	60	58	43	37	23

Table 7.15: Number of days with only one to four values which are not included in the CA.

Threshold for an Avalanche Day	Northern Cairngorms			Creag Meagaidh			Total		
	Threshold Missing Values	Nr of Deleted Days		Threshold Missing Values	Nr of Deleted Days		Threshold Missing Values	Nr of Deleted Days	
		[abs]	[%]		[abs]	[%]		[abs]	[%]
>=1	>= 8	24	5.6	>= 10	4	1.3	>= 7	565	12.3
>=2	>= 9	3	2.1	>= 10	3	1.9	>= 7	361	12.7
>=3	-	-	-	-	-	-	>= 8	99	5.2
>=4	-	-	-	-	-	-	>= 8	68	5.2
>=5	-	-	-	-	-	-	>= 8	52	5.6
>=6							>= 8	45	7.4
>=7							>= 8	34	7.7
>=8							>= 9	1	0.3
>=9							>= 9	1	0.4
>=10							-	-	-

Outliers

Outliers in the data can influence the results of a CA heavily (e.g. Mooi & Sarstedt 2011, Pang-Ning *et al.* 2006). Therefore outliers are examined and they belong in most cases to days showing extreme values, especially within the variable 'Summit Air Temperature Difference'. But as these values are still realistic and typical especially for Scotland, they are included in the analysis. Table 7.16 shows for different thresholds of an avalanche day the dates which would have acted as outliers.

Table 7.16: Outliers for each area considering different thresholds for an avalanche day.

Definition of Avalanche Day	Northern Cairngorms		Lochaber		Glencoe		Southern Cairngorms		Creag Meagaidh	
	Dates	Nr	Dates	Nr	Dates	Nr	Dates	Nr	Dates	Nr
>=1	22.03.13, 22.02.97, 16.3.11, 11.3.02, 24.1.08	**1**/5	13.01.13, 21.12.11	**1**/2	24.2.10, 4.2.12, 15.1.99, 18.1.98	**1**/3	28.2.2012, 16.3.11	2	7.2.12, 16.1.99, 23.12.98, 17.1.99, 18.2.00, 23.12.09, 25.12.09	2/**3**/7
>=2	11.03.02, 16.03.11	2	21.12.11, 16.1.93, 14.2.13, 4.2.12, 17.1.98, 30.1.09	**1**/4/6	8.2.96, 25.2.96, 24.2.10, 16.1.00	**1**/4	26.2.2010, 28.2.00	2	16.1.99, 17.1.99, 23.12.98, 28.1.12	4
>=3	11.03.02, 16.03.11	2	16.1.93, 14.2.13, 4.2.12, 16.2.00	3	16.01.00	1		0	16.1.99, 17.1.99, 4.2.03	1/**3**
>=4	16.03.11, 11.03.02	1	14.2.13, 16.1.93	0	16.1.00, 15.3.10, 9.2.94, 5.1.00	0		0	16.1.99, 17.1.99, 7.3.97, 4.2.10	**1**/4
>=5	05.04.10	1	04.02.12	1	01.03.10	1	not possible		16.01.99	1

7.3.2.2 Method

As main clustering algorithm a hierarchical clustering is chosen, for several reasons. Dates with missing values can generally be included in the analysis. The possibility to plot a dendrogram gives a great view on each step that is going on in the clustering process and no black box algorithm introduces even more uncertainties.

The main analysis is done using Ward's method as it is often proposed in literature (e.g. Backhaus *et al.* 2008, Mooi & Sarstedt 2011, Burns & Burns 2008).

7.3.2.3 Number of Clusters

Screeplots (see figure 7.16), number of events, dendrogram analysis and comparisons between different scenarios (see appendix A.9 for detailed plots) are used to decide on the number of clusters that are processed for each area. In the end, one scenario that represents the other ones best and can well be described and interpreted, is chosen for every area and for the total amount of data. The influence of this choice on the results is analysed and other scenarios are taken into account when interpreting.

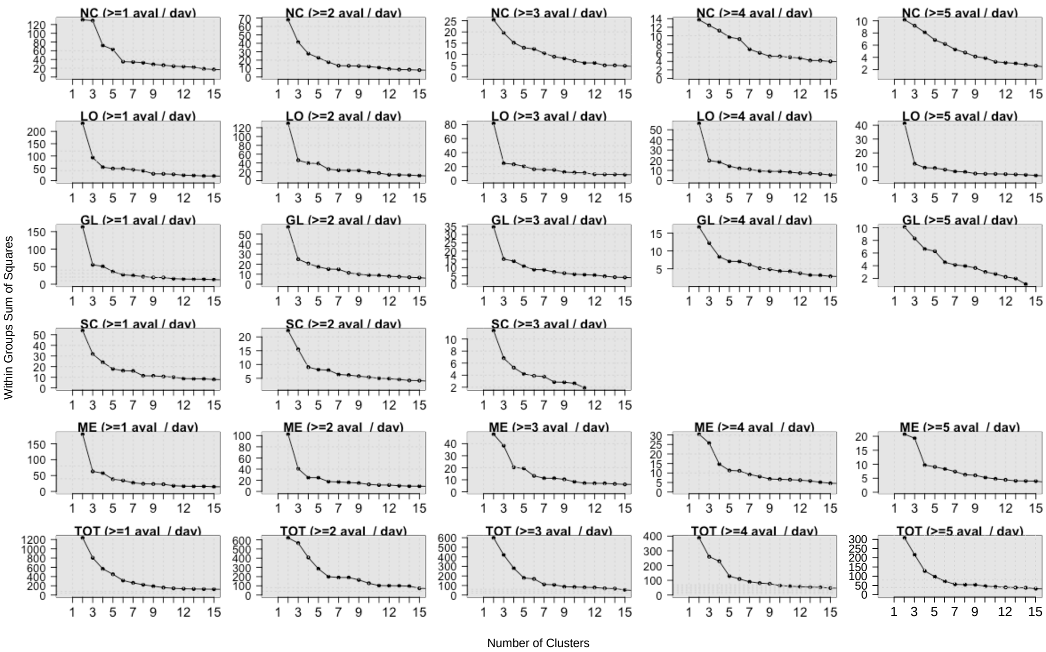


Figure 7.16: Screeplots for all five SAIS areas and the total amount of data considering different thresholds for an avalanche day.

7.3.2.4 Resulting Patterns

One optimal scenario for each area and the total amount of data is summarised in tables 7.17 to 7.22. The threshold for an avalanche day is chosen such that the other scenarios are best represented. The number of clusters is decided depending on the intensity of their characteristics and their differentiability. Generally, the more clusters are built, the more pronounced their characteristics get. But with a too high number of clusters, they tend to become too similar and no additional information can be provided.

Some very distinct and strongly characterised clusters are found, which occur in the vast majority of scenarios. They can be seen as major patterns concerning Scottish avalanches[4] (see table 7.23 for an overview of the cluster means).

- One obvious cluster is characterised by warm summit air temperatures, with means clearly above the melting point (without exception between 0 °C and +5 °C) and strong short time warming with averages of +3 °C to +6 °C. These summit air temperatures are measured at elevations which are higher than the majority of avalanche releases. Therefore it is likely that temperatures at main avalanche altitudes are even warmer, provided that no inversions are present.
  With rising air temperatures, snow temperatures tend to increase as well. The snow cover can be influenced by liquid water, gets wet, destabilises and a higher avalanche activity results, including many natural releases. These characteristics are confirmed in the clustering results when looking at the comment information. It is the only cluster in the majority of scenarios, in which clearly more avalanches are classified as wet than dry (e.g. TOT: 207 versus 90, ME: 29/77 versus 12/44[5] or LO: 82/71 versus 20/33). Further, little to no avalanches are reported to be human triggered (e.g. TOT: 27 versus 166 (naturally), GL: 11 versus 34 or NC: 19 versus 28), and a considerable amount of avalanches is mentioned to be related to cornice collapses. Warm conditions with heavy, wet snow can destabilise the intrinsically unstable cornice structure considerably in short time.
  Also the distribution throughout the months confirms warm conditions, as a major part of the avalanches releases in March and April and during the day between 9 am and 3 pm. However, the presence of high frequencies also in December and January emphasises the fact that warm conditions can arise throughout the whole Scottish winter and are not only confined to the typical spring months, as it would be the case for example in the Alps. This has already been stated by Ward *et al.* (1985), but he added that thaws tend to be short and marginal in January and early February.
  An additional characteristic of the cluster is noticeable low mean values of the Weakness Index at about 4 or 5 (maximum 5.8 in the Northern Cairngorms), which are clearly lowest in comparison to all the other clusters. Weak crystals grow primarily during cold periods and sintering processes take preferred place in warming conditions, snow grains get larger, rounded, well bonded and the stability increases. The sudden instability that arises when temperatures get very close to the melting point or due to the influence of liquid water, is not systematically included in the snow profile data and is therefore not considered in the Weakness Index.
  The cluster is represented clearly in all areas but Glencoe. Glencoe indeed shows short time warming in each of its clusters, but in no one warm air temperatures are present, as they remain without exception below 0 °C. In all other areas, the characteristics of the cluster are neither dependent on the number of clusters, nor on the threshold that is used for an avalanche day. With higher thresholds for an avalanche day the characteristics tend to be even more pronounced (especially in Creag Meagaidh and Lochaber). These are indications that warm conditions play a very important role for the avalanche activity during the Scottish winter.

  ⇒ ***Pattern 1: Thawing situation***

[4]General interpretations are not only based on values from tables 7.17 to 7.22, but especially when speaking about general variable values, on the plots showing a broader range of information in appendix A.9.

[5]The first values represents the optimal solution summarised in tables 7.17 to 7.22, the second ones those with thresholds >=1 avalanches/day.

Table 7.17: Summary of the results from the CA for the Northern Cairngorms (>=3 avalanches per day / 5 clusters).

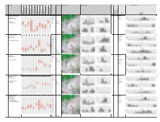


Table 7.18: Summary of the results from the CA for Lochaber (>=3 avalanches per day / 4 clusters).

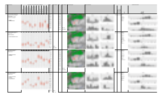


Table 7.19: Summary of the results from the CA for Glencoe (>=3 avalanches per day / 4 clusters).

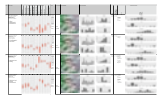


Table 7.20: Summary of the results from the CA for the Southern Cairngorms (>=2 avalanches per day / 5 clusters).

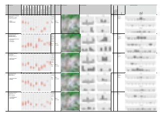


Table 7.21: Summary of the results from the CA for Creag Meagaidh (>=3 avalanches per day / 6 clusters).

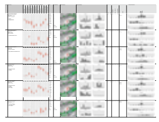


Table 7.22: Summary of the results from the CA for the total amount of data for all SAIS areas (>=3 avalanches per day / 5 clusters).

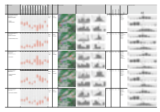


Table 7.23: Overview of cluster means for the optimal scenarios (tables 7.17 to 7.22) for different areas and the total amount of data.

Pattern	Area	Summit Air Temperature [°C]			Summit Wind Speed [mph]			Snow Index			Summit Air Temperature Difference [°C]	Weakness Index
		Avalanche Day	3 Day Sum	7 Day Sum	Avalanche Day	3 Day Sum	7 Day Sum	Avalanche Day	3 Day Sum	7 Day Sum		
Thaws	**NC**	1.7	-5.1	-9.0	41.0	51.2	143.6	4.7	14.7	28.0	5.4	5.8
	LO	0.8	-3.7	-10.1	32.7	82.3	177.0	0.5	9.8	25.5	3.3	4.2
	SC	2.5	1.4	-8.6	20.3	62.1	155.4	0.4	1.7	15.0	2.6	5.2
	ME	2.3	-2.5	-18.9	34.0	63.8	136.2	0.1	8.1	31.1	4.1	4.5
	Total	0.6	-4.4	-13.8	31.9	80.3	173.9	1.8	8.5	22.5	3.0	5.0
	Mean	**1.6**	**-2.8**	**-12.1**	**32.0**	**67.9**	**157.2**	**1.5**	**8.5**	**24.4**	**3.7**	**4.9**
Cold Periods	**NC**	-4.7	-15.4	-36.8	33.2	76.8	187.2	5.1	5.4	10.7	0.4	6.0
	LO	-5.0	-13.6	-26.8	19.6	69.6	166.9	1.2	8.5	17.4	-0.2	5.5
	SC	-3.9	-11.7	-21.5	36.2	95.7	164.8	6.4	19.8	31.3	-0.4	7.1
	ME 1	-6.7	-5.8	-5.3	11.4	32.9	65.7	5.3	9.4	16.0	-4.5	6.6
	ME 2	-4.6	-16.5	-32.8	18.5	47.3	113.4	6.2	18.7	33.4	1.8	7.6
	Total	-6.2	-19.1	-38.7	20.2	59.2	142.0	4.6	11.1	23.3	0.4	6.6
	Mean	**-5.2**	**-13.7**	**-27.0**	**23.2**	**63.6**	**140.0**	**4.8**	**12.2**	**22.0**	**-0.4**	**6.6**
Drifting and New Snow	**NC**	-2.2	-9.2	-23.9	45.0	137.2	318.4	6.3	17.1	41.1	1.2	6.0
	LO	-2.4	-8.1	-15.7	26.5	82.5	186.5	7.4	18.3	33.9	0.7	5.1
	GL	-1.1	-8.4	-26.4	33.9	79.4	176.0	2.2	13.2	34.8	3.1	5.7
	SC	1.4	-2.0	-12.3	33.3	107.7	222.3	0.5	13.5	24.5	2.7	5.0
	ME	-4.3	-11.1	-22.8	22.5	81.9	186.4	8.0	24.1	48.6	-0.3	5.9
	Total 1	-2.8	-7.1	-13.8	24.9	74.1	167.0	6.8	16.8	32.5	-0.4	5.3
	Total 2	-3.6	-9.5	-20.4	36.8	109.1	239.5	4.8	11.5	23.9	-0.4	6.2
	Mean	**-2.1**	**-7.9**	**-19.3**	**31.8**	**96.0**	**213.7**	**5.1**	**16.4**	**34.2**	**0.9**	**5.6**
Not extreme	**NC**	-3.8	-11.6	-26.7	37.7	114.1	270.4	4.9	12.0	23.7	0.3	6.1
	GL	-0.3	-2.6	-9.1	24.0	64.3	141.5	2.7	5.3	21.6	1.0	5.7
	SC	-3.2	-8.0	-17.1	29.3	61.4	152.9	6.3	9.3	15.3	-0.8	5.7
	ME	-4.6	-16.5	-32.8	18.5	47.3	113.4	6.2	18.7	33.4	1.8	7.6
	Total	-3.0	-9.8	-21.4	19.0	56.5	137.9	1.0	6.6	18.3	0.4	5.5
	Mean	**-3.0**	**-9.7**	**-21.4**	**25.7**	**68.7**	**163.2**	**4.2**	**10.4**	**22.5**	**0.5**	**6.1**

- The second well pronounced cluster shows very cold air temperatures, not only at the avalanche day, but also in the three and seven day sum. Means stay mostly below -5 °C and are around -15 °C in the three day and -35 °C in the seven day sum. Additionally, the avalanche days show highest values of the Weakness Index with averages of at least 6 and up to 8 in the majority of scenarios. Long cold periods lead to the development of weak crystals, which destabilise the snow cover, promote natural avalanche releases but are also prone to triggering. Indeed, the majority of avalanches are reported to have released naturally, but also considerably more human triggered avalanches are present than in the thawing cluster, in some cases almost about 50% (Lochaber, Northern Cairngorms, Total). Some avalanches are again mentioned to be related to cornices, probably due to collapses after the development of weak layers. Little to almost no avalanches are reported to be wet instead of dry (e.g. TOT: 6 versus 91, SC: 0 versus 5/9, ME: 5/9 versus 22/70, LO: 9 versus 17, NC: 5/1 versus 37/25 and 0/5 versus 9/67). Maximum avalanche frequencies can be seen in February and sometimes January, when also lowest temperatures in Scotland exist, but also avalanches which occurred in March are assigned to this cluster.

  A special characteristic of the cold avalanches are maximum frequencies of very low gradients (< 20°). This fact has already been stated by Diggins (2010, 1): "natural avalanche releases on slopes less than 15° were noted during field tests in this region and the cause for failure was often, depth hoar or buried surface hoar". Additionally, the spatial distribution of the avalanches is in most areas not so strongly confined to the main avalanching areas and a larger amount of spatial outliers occurs.

  These can be indications that the development of weak layers is not constrained to steeper slopes, but rather can influence the stability of the snow cover at a larger extent. The cluster is represented well in the majority of scenarios. In all scenarios including those with few clusters the summit air temperatures can be clearly distinguished from the other ones, except in the solutions for Glencoe. As summit air temperatures do not show large differences in this area, the cold cluster is not clearly pronounced. In general, this cluster seems to be very important for the Scottish winter and can lead to unexpected hazards in many situations. Conditions are less obvious and sometimes also less predictable than other situations. This is especially critical, as cold periods are often characterised by clear skies and no rain, which are best conditions for recreation. On such days, more than an average amount of people is probably out in the hills which increases the avalanche risk automatically (Moss 2009, Ward 1980).

  Wind speeds and snow indices are pronounced differently in various areas and scenarios. Generally, summit wind speeds show values of about 20 mph, 70 mph and 150 mph[6] and a snow index around 0 to 5, 5 to 10 and 10 to 20. In the Northern Cairngorms and Creag Meagaidh, a cluster is present additionally, which is characterised by higher amounts of new snow (values of about 8, 20 and 40). In the Northern Cairngorms, wind speeds with values of 40 mph, 70 mph and 200 mph are higher than in the other areas. In the Southern Cairngorms, the cold cluster with minimum temperatures shows at the same time maximum values of wind speeds and snow indices.

  ⇒ ***Pattern 2: Cold periods***

- The third very distinct cluster includes days with a high amount of new snow and very high wind speeds, at the avalanche day as well as in the three day and the seven day sum. Summit wind speeds show average values of at least 20 mph, around 80 mph to 100 mph in the three day sum and up to 180 mph in the seven day sum (in the Cairngorms even 300 mph). Snow index means lie at around 15 to 20 in the three day sum and reach 30 or 40 in the seven day sum.

  Much new snow and high wind speeds lead to drifting with the accumulation of dangerous windslabs that are prone to avalanching. Again, some cornice avalanches are mentioned,

[6] The combination of these three values always represents the mean at the avalanche day, the three day sum and the seven day sum.

which is not surprising as cornices are results of drifting processes at ridges and consist of windslab snow.
A considerable amount of avalanches are reported to be human triggered. Avalanche frequencies throughout the week show maximum values at Wednesdays, Saturdays and Sundays (Total, Lochaber, Northern Cairngorms and Creag Meagaidh), which correlates with the days at which probably most people are out in the hills. Most avalanches release during the main avalanche season between January and March.
The cluster is very well pronounced in most areas and scenarios, and is in almost any case related to mean air temperatures between 0 °C and -5 °C, about -10 °C and -20 °C. Only for the Southern Cairngorms the temperatures are lower, as the cluster represents on the same time the cold pattern. Summit wind speeds are in all areas comparable with values around 30 mph, 100 mph and 200 mph, but reach means of 40 mph, 150 mph and 300 mph in the Northern Cairngorms.
Cluster sizes seem to be rather on the large side (maximal in Glencoe, Lochaber and Total), which emphasises the importance of drifting processes for Scotland. It has already been mentioned by Barton & Wright (2000, 58) that many of the most destructive avalanches in Scotland have followed "a week or more of consistently cold weather with snow showers or drifting". In the areas Lochaber, Glencoe, Southern Cairngorms and Creag Meagaidh, two distinct drifting clusters result, with one being more pronounced than the other.

Table 7.24: Beaufort Scale (MetOffice 2013c).

Beaufort Scale	Description	Average Wind Speed [mph]	Beaufort Scale	Description	Average Wind Speed [mph]
1	Light air	1.2-3.0	7	Near gale	32.0-38.0
2	Light breeze	3.7-7.5	8	Gale	39.0-46.0
3	Gentle breeze	8.0-12.5	9	Severe gale	47.0-55.0
4	Moderate breeze	13.0-18.6	10	Storm	56.0-64.0
5	Fresh breeze	19.3-25.0	11	Violent storm	65.0-74.0
6	Strong breeze	25.5-31.0	12	Hurricane	75+

Depending on the thresholds, one can also assume that the extreme drifting clusters should rather be named 'Storm'. It is difficult to estimate the detailed amount of new snow that falls during a period, as only the subjective snow index values are available. However, values above 4 or 6 can be seen as considerable amounts of new snow and 8 even means continuous snow over a longer period (see table 4.3). When considering the Beaufort Scale describing the intensity of wind speeds (see table 7.24), a storm is defined with having average wind speeds of about 55 mph. Such high wind speeds are with no cluster means reached.

⇒ ***Pattern 3: Drifting conditions with much new snow***

- Lastly, almost in every scenario a cluster characterised by not such extreme conditions is present. Depending on the data, the avalanche days show some wind (means of about 20 mph, 50 mph and 100 mph to 150 mph, little higher in the Northern Cairngorms with 30 mph, 70 mph and 200 mph), some new snow (means of 2, 5 to 10 and 15 to 20), are relatively cold without showing extreme temperatures (about -3 °C, -5 °C to -10 °C and -10 °C to -20 °C) and Weakness Indices lie at about 5. This cluster may include those situations which are most dangerous, as no obvious pattern is present and conditions include no clear sign for high avalanche hazards. Therefore it is little surprising, that a considerable amount of avalanches are reported to be human triggered.

  ⇒ ***Pattern 4: Not extreme conditions***

Additionally to the four patterns, the analysis has also shown some more general characteristics.

- Strong short time cooling with high negative values of the air temperature difference are found in no scenario. Means in clusters belonging not to the thawing pattern are always very close to 0 °C.
- Cornices are mentioned in almost each cluster for every area and scenario, which lets assume that they are dangerous features in general and are not highly dependent on the special pattern that is present.
- Some differences between the areas become apparent. Most obvious are the considerably higher wind speeds in the Northern Cairngorms, probably due to the fact that the summit measures are taken at Cairn Gorm which is more than 100 meters higher located than the stations at Aonach Mor or Cairnwell. McClatchey (1996, 37) mentions also that "the Cairn Gorm summit AWS is extremely exposed and strong winds are accelerated over the relatively smooth rounded summit" The cold cluster in the Southern Cairngorms is the only one in which also maximum snow indices and wind speeds are present. For Glencoe least obvious patterns are found. Air temperatures are similar in all clusters and also wind speeds show no large differences. Therefore only very high snow index values in all three measures of cluster 3 can be well distinguished. Wind speeds are more similar in the cluster for Lochaber than in other areas and show no means above 40 mph, 80 mph and 200 mph. "Because the Cairngorms are further from the sea than other climbing areas and many of the cliffs are very high, the conditions here do not fluctuate as rapidly as elsewhere" (Fyffe 2000, 8).
- The terrain analysis of the avalanches does not show large differences between the clusters. In Creag Meagaidh nearly no avalanches release on NW slopes, but more on SW to S slopes than in the other areas. Maximum altitudinal frequencies differ slightly between the areas, with a peak at 1200 masl for Lochaber, 800 masl to 900 masl for Glencoe, 900 masl to 1000 masl for the Southern Cairngorms and Creag Meagaidh, and 1000 to 1100 masl for the Northern Cairngorms.
- The activities related to the avalanches provide only little important information, especially for Creag Meagaidh and Southern Cairngorms, where almost no information is available. Considering the scenario for the total amount of data and the Northern Cairngorms, by far most avalanches are related to climbing, some less to skiing, and a considerable amount is reported by observers or the snow patrol. Nevertheless, differences between the clusters cannot be seen.
- The distributions throughout the seasons seem not to show general patterns, and also cluster sizes or the average number of avalanches per day provide no interesting information.
- The characteristics of single clusters can better be distinguished using the total amount of data. The dependency on the amount of input data is much smaller than for the separate areas and the characteristics of the clusters are more pronounced. However, the same clusters can also be found using much less input data for the separate areas. Differences between the thresholds used for an avalanche day are smaller when using the total amount of data, but quantile ranges in general tend to be larger.
- Scenarios using higher thresholds for an avalanche day are tested for the whole amount of data (see figures A.33 and A.34). Using thresholds between 6 and 10, the patterns are still identical and get even more pronounced as medians come closer together and quantile ranges get smaller. With thresholds above 10, the amount of input data becomes too limited to ensure reliable results.

Patterns of Non-Avalanche Days

To enable a comparison of the scenarios to conditions on non-avalanche days, also a clustering of the days on which no avalanches occurred is done (see figure 7.17). The following differences can be seen:

- Minimum air temperatures are comparable in avalanche and non-avalanche clusters. Medians which reach low values of about -7 °C to -8 °C in avalanche clusters stay above -5 °C. Biggest differences show maximum air temperatures since they lie in avalanche clusters at 5 °C but reach even more than 10 °C in non-avalanche clusters. Some seldom outliers in the 'thawing' clusters of avalanche days reach temperatures of 15 °C in the three day sum and 20 °C in the seven day sum, in non-avalanche clusters are the respective values with 30 °C and 60 °C multiplied.

- Wind speeds show no important differences between avalanche and non-avalanche clusters, neither in medians, nor in minimum or maximum values.
- Amounts of new snow differ highly and are considerably lower in non-avalanche clusters than in avalanche clusters, mainly in the three day and seven day sum. Medians in non-avalanche clusters are at maximally 13 and 20 in comparison to 20 and 40 in avalanche clusters. Further, minimum values in non-avalanche clusters lie without exception at 0 and values of 4 (one day), 20 (three day sum) and 30 (seven day sum) are reached in avalanche clusters.
- Air temperature differences are comparable.
- The Weakness Index does not show extreme differences, but minimum values in avalanche clusters lie only very seldom below 3 in comparison to always 0 in non-avalanche clusters. Maximum values are comparable, and medians at 5 to 7 tend to be higher in avalanche clusters than in non-avalanche clusters (values of 4 to 6).

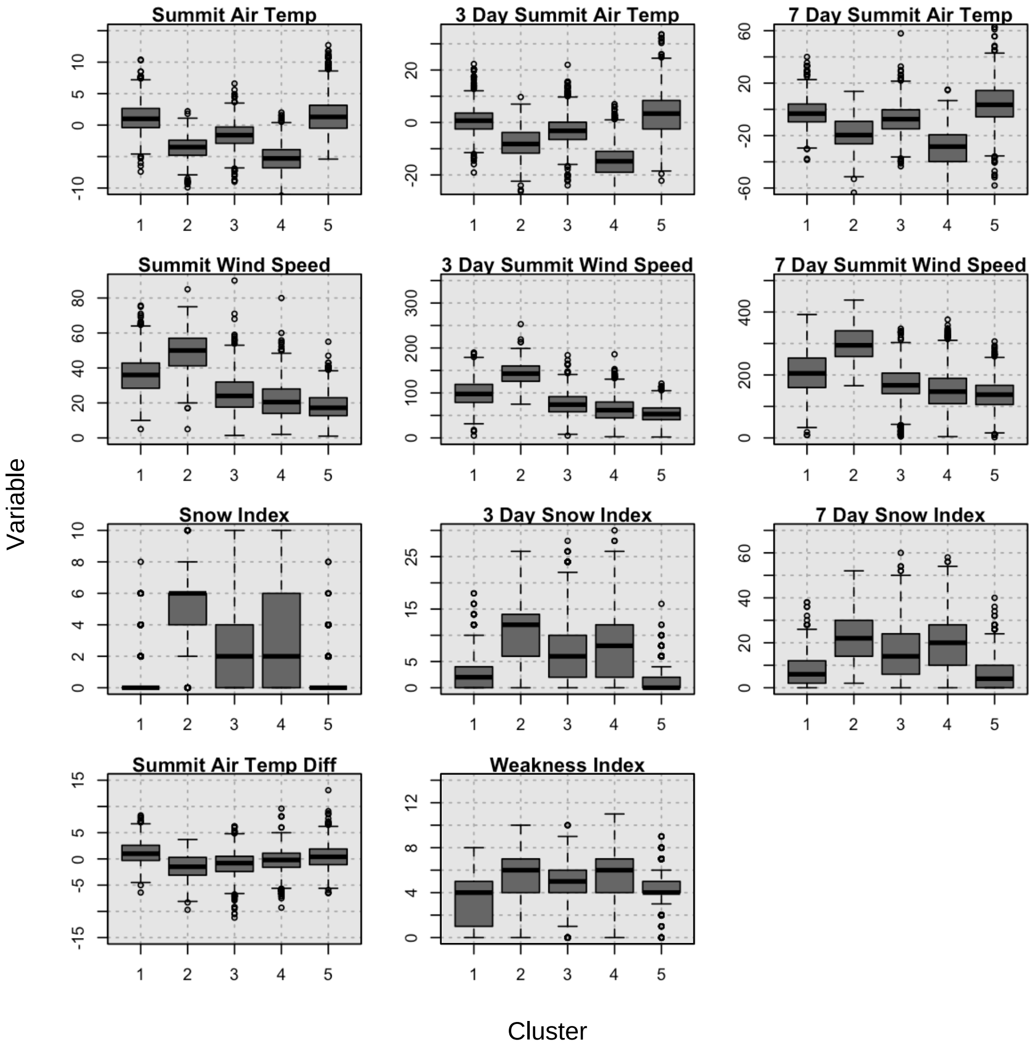


Figure 7.17: Clustering solution of non-avalanche days.

7.3.2.5 Validation

The validation of the CA is done on a variety of aspects. Every decision on parameters and every choice of method can have an effect on the results. Therefore the influences of the following parameters on the resulting patterns are examined:

Input Data
Different thresholds of an avalanche day are considered in each area as always five scenarios (>=1 to >=5 avalanches per day) are compared (see appendix A.9 for detailed plots). Generally, the scenarios based different thresholds do not show considerable differences, at least for thresholds >=1, >=2 and >=3. The higher the threshold, the fewer days are included in the input data and the more are the characteristics of each cluster dependent on a single day. Therefore the differences in the >=5 avalanches/day scenarios are biggest in comparison to the others. Cluster sizes decrease sometimes to 1 or 2 and these solutions are not reliable any more.

Generally, the quantile ranges in the boxplots become smaller the higher the threshold is chosen. Therefore the fewer days are included or the higher the avalanche activity on each day, the more pronounced the characteristics of the patterns seem to be. The clearer a pattern is visible, the higher the number of avalanches which has to be expected gets.

Summarising, the choice of the threshold (or the pronouncement of avalanche activity) does not have much influence on the general patterns, but determines the intensity of their characteristics.

Outliers and Omitting Days
The differences between the hierarchical clustering using the entire data set, the data without outliers ('26 December 2011' and '23 March 2013' are chosen as most important outliers) and the data where the days with missing values are omitted are shown in figure 7.19 for the solution of the total amount of data with 5 clusters and a threshold of >=2 avalanches/day.

When outliers are excluded from the CA, medians are by trend less extreme, but as only little outliers are present in the data, the influence is marginal and the patterns stay the same. A considerable amount of information is not included in the analysis when days with missing data are omitted (see table 7.15). However, differences are still small and the general patterns do not change.

The influence of outliers is probably larger when using less input data in the CA for the separate SAIS areas, but also there, the resulting patterns will not change.

Clustering Algorithm
The influence of the hierarchical clustering method is examined by comparing the results with those from the application of a k-means algorithm (see figure 7.19). As the k-means can only be run on data without any missing values, the results have to be combined to a hierarchical clustering applied on the same input data in which every day with missing values is deleted. This leads for the scenario using a threshold of >=2 avalanches/day in an CA with 5 clusters for the total amount of data to the use of 1524 days out of totally 3826. A k-means algorithm does not always give the same results when conducted several times, but applying the algorithm in an iteration shows that about 5 different stable solutions exist. The one with cluster sizes 238, 329, 190, 389, 380 is most frequent with a betweenSumOfSquares/totalSumOfSquares of 40%.

Even though the fact is confirmed that k-means tends to produce clusters of similar sizes (e.g. Backhaus *et al.* 2008, Steinley 2006), the resulting patterns are very similar and the choice of the clustering algorithm has therefore only little influence on the resulting patterns.

Order of Input Data
Several authors mention that the order of the input data influences the clustering result (e.g. Backhaus *et al.* 2008, Mooi & Sarstedt 2011, Pang-Ning *et al.* 2006). This effect is exemplarily tested and three versions for the total amount of data are shown in figure 7.20, each with randomly ordered rows in the input data.

The results are slightly different, but in no cluster and no variable considerable differences are present. The patterns stay the same and the order of the input data influences the resulting patterns not significantly.

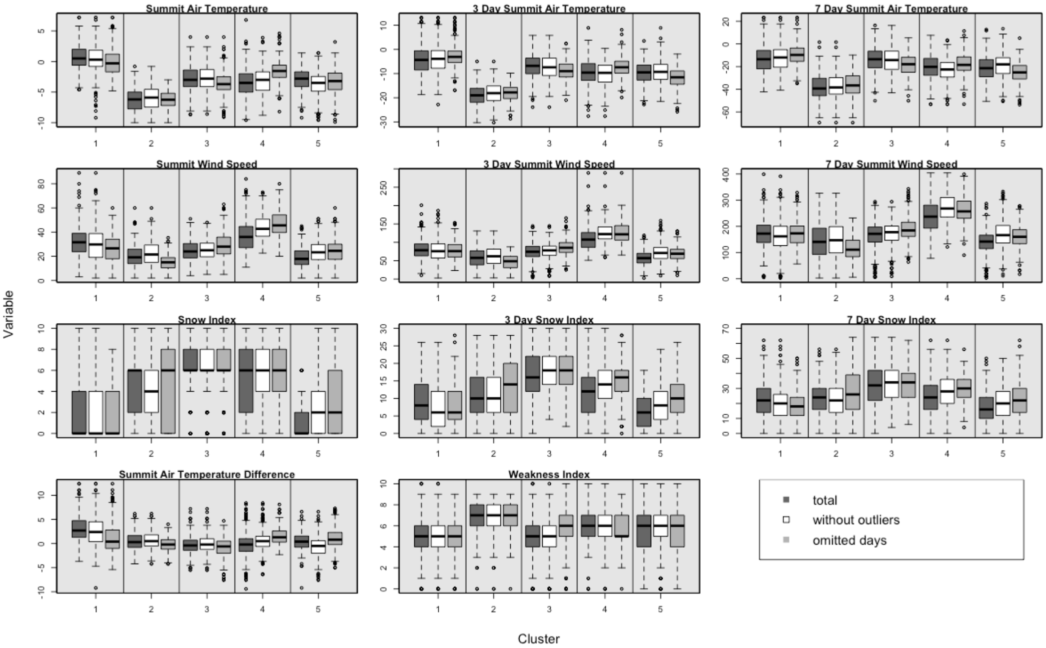


Figure 7.18: Comparison of a hierarchical clustering applied on the total amount of data, on data without any outliers and on data where days with missing values are omitted (Cluster sizes: total: 528, 251, 678, 614, 401, without outliers: 602, 291, 640, 311, 626, with omitting days: 437, 128, 401, 134, 426).

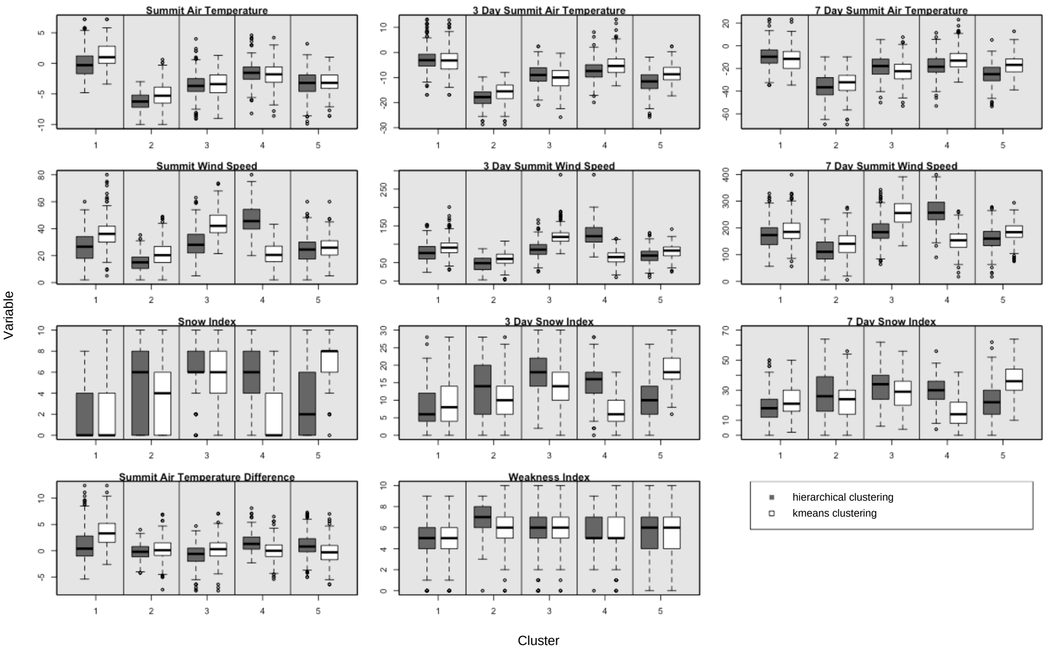


Figure 7.19: Comparison of the clustering solutions using either a hierarchical or a k-means algorithm (Cluster sizes: k-means: 238, 329, 190, 389, 380 | hierarchical clustering: 437, 128, 401, 134, 426).

Amount of Input Data

Some authors propose as validation method of a clustering result to run the cluster analysis only on half of the input data (e.g. Mooi & Sarstedt 2011). As a similar approach is already done by running the clustering algorithm on different data sets for each area as well as for the total amount of data, this is not done additionally.

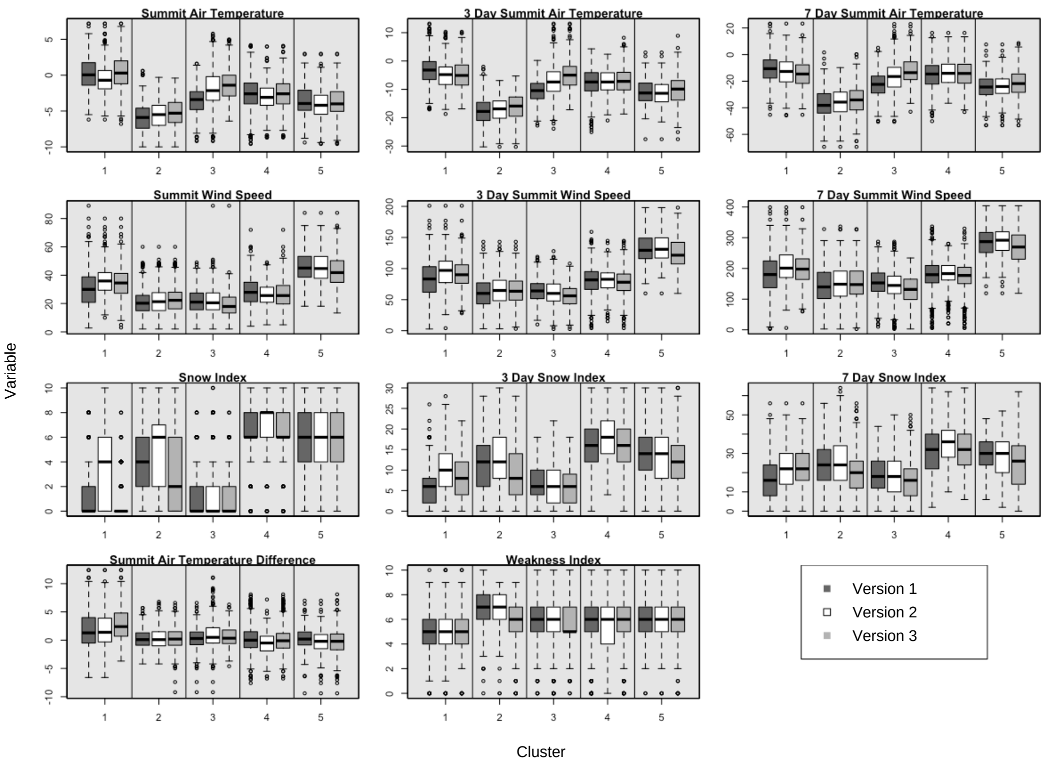


Figure 7.20: Comparison of three clustering solutions with randomly ordered rows in the input data matrix.

7.3.3 Principal Component Analysis

As the PCA can only deal with days on which no data is missing, the amount of input data decreases considerably with each variable that is considered (see table 7.25). In the majority of solutions, more than half of the input data cannot be considered which makes the application of the PCA as support for the decision on the input variables for the CA not sensible. Therefore on the input variables for the CA is decided without considering the PCA. However, a PCA is applied to assess the quality the used input variables.

When looking at the PCA for the scenario of the total amount of data, it can be seen that the first 5 components explain more than 80% of the total variance and that each component is clearly correlated with specific input variables (see table 7.26). The first component which explains about one fifth of the total variability is determined highly by the summit air temperature measures. The second component explains only little less of the total variability and the snow index variables are most important, especially the three day sum. The third component is correlated to the wind speed, the fourth to the air temperature difference and the fifth to the Weakness Index. Including the fourth and fifth component rises the explained total variability each about 10% from 63% to 75% and 84%.

This result is little lower but comparable to other studies in which a PCA has been applied to similar avalanche input data (see section 2.3.2). Birkeland *et al.* (2001) could explain about 93% of the total variability with five components, Esteban *et al.* (2005b) 90% and García *et al.* (2008) 94% using both six components and Mock & Kay (1992) could explain 85% using three components.

In Birkeland *et al.* (2001) and Mock & Kay (1992), the first components are maximally loaded by the two variables snow water equivalent and amount of new snow fall, whereas the other authors relate the components not to single variables but to atmospheric situations. The snow water equivalent is not included in the current data and could therefore not be considered in the analysis, but the snow index shows similar importance as found by Birkeland *et al.* (2001) or Mock & Kay (1992). Instead, the air temperature seems to play a very important role in the current analysis, which has not been mentioned by any of the other authors.

Table 7.25: Number of input entries which would be available when including different combinations of variables in the PCA (total amount of data: 11123).

Air Temperature	Wind Speed	Wind Direction	Cloud	Drift	Foot Pen	Ski Pen	Total Snow Depth	Summit Air Temperature	Summit Wind Speed	Summit Wind Direction	Rain	Snow Index	Number of Entries when Omitting Days with Missing Values [abs]	[%]
✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	976	8.8
			✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	1007	9.1
✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	1158	10.4
✔	✔	✔	✔	✔	✔		✔	✔	✔	✔	✔	✔	4441	39.9
✔	✔	✔	✔	✔			✔	✔	✔	✔	✔	✔	4462	40.1
			✔	✔			✔	✔	✔	✔	✔	✔	4543	40.8
				✔			✔	✔	✔	✔	✔	✔	4570	41.1
				✔			✔	✔	✔	✔		✔	5211	46.8
							✔	✔	✔	✔		✔	5212	46.9
								✔	✔	✔		✔	7785	70.0

Table 7.26: Results of the PCA conducted on the total amount of days for all five SAIS areas on which at least 2 avalanches occurred (grey: highest correlations).

Component --- Variable |	PC1	PC2	PC3	PC4	PC5	PC6
Summit Air Temperature	**0.48**	-0.20	-0.16	0.29	0.14	-0.03
3 Day Summit Air Temperature	**0.47**	-0.12	-0.37	-0.17	0.03	-0.01
7 Day Summit Air Temperature	**0.39**	-0.02	-0.46	-0.27	0.05	0.03
Summit Wind Speed	0.33	0.23	**0.41**	0.11	0.14	0.30
3 Day Summit Wind Speed	0.32	0.38	**0.35**	-0.21	0.00	-0.04
7 Day Summit Wind Speed	0.29	0.37	**0.25**	-0.26	-0.09	-0.36
Snow Index	-0.11	**0.45**	-0.28	0.05	0.28	0.53
3 Day Snow Index	-0.08	**0.50**	-0.28	0.24	0.01	0.11
7 Day Snow Index	-0.04	**0.37**	-0.28	0.40	-0.15	-0.60
Summit Air Temperature Difference	0.22	-0.16	0.21	**0.66**	0.26	-0.01
Weakness Index	-0.17	0.00	0.01	-0.21	**0.88**	-0.35
Standard Deviation	1.72	1.57	1.24	1.13	0.99	0.79
Proportion of Variance	0.27	0.23	0.14	0.12	0.09	0.06
Cumulative Proportion	0.27	0.49	0.63	0.75	**0.84**	0.89
Kaiser Kriteria	2.95	2.48	1.53	1.27	0.99	0.63

8

DISCUSSION

The results presented in chapter 7 are discussed and the more remarkable findings are analysed in a broader context. In the end of each section, the research questions proposed in section 1.2 and chapter 5 are answered. Referring to part A, mainly comparisons to Ward (1984a) are drawn, because it is the only other survey of avalanche activity in Scotland. For part B, primarily the research questions are answered and the most important aspects of the combination of the different data sources and the resulting data set are summarised. The derived patterns resulting from part C are further interpreted and compared to other studies.

8.1 Part A: Overview of Avalanche Activity

Spatial Distribution, Spreading, Human Activity and Bias

The general distribution of the total amount of avalanches can be used to analyse the three main characteristics concerning the locations of avalanches stated by Ward (1984a, 95) about 30 years ago:

1) "Avalanches are very widespread."
2) "Certain locations appear to produce more avalanches than others."
3) "Some slopes appear to avalanche more often than others."

These three statements can generally be confirmed with the current results, but some additional factors have to be considered.

Avalanches are widespread, but they show a clear spatial distribution. The vast majority releases in the Scottish Highlands, as avalanche occurrences are confined to a certain terrain. This can already be seen when looking at the distribution of gradients and aspects in figure A.2.

Ward (1984a) as well as the current avalanche data show that most avalanches release on slopes which are steeper than 30° and that most avalanche prone slopes have gradients between 35° and 45°. However, under special conditions with long and very cold temperatures and the development of widespread depth hoar, also avalanche releases on slopes with gradients smaller than 15° are possible (Diggins 2010). The analysis of the DEM in this thesis shows similar results. About 20% of the avalanches occurred on slopes of less than 30° and the highest frequency is on slopes between 30° and 35°. As the DEM resolution with 50 meters is relatively high, the slopes are usually underestimated due to smoothing effects of the topography (Kienzle 2004, Thompson *et al.* 2001, Zhang *et al.* 1999). Therefore the results are well comparable to Ward (1984a) and the information included in the current data set of the reported avalanches.

The dominance of northeastern aspects (Ward 1984a) can be confirmed with the actual data as well, adding Northerly oriented slopes in general as avalanche prone aspects (see figure 8.2). Ward (1984a) explains this fact firstly with the orientation of the main climbing slopes and secondly with meteorological factors such as snow drift. Indeed, wind directions included in the snow profile data show the highest frequencies at southwest (second highest at south and west, see figure 8.1) and also Buchan (1890, xl) mentions northwestern to eastern winds as least frequent ones. Therefore drifting snow is preferably accumulated on Northern aspects leading to an increased avalanche activity at those aspects.

The distribution of total altitudes, aspects and gradients in the bounding boxes of the SAIS areas and the amount of avalanches that occur within the corresponding ranges are shown in figure 8.2. On the one hand, differences between the five areas can be seen, such as the large amount of high altitudes in the Cairngorms or the high frequency of relatively steep slopes between 30° and 40° in Glencoe. On the other hand, those parts of the available terrain on which avalanches occur get apparent. Releases take only in the small upper part of altitudes and gradients place and no upper boundaries of maximum values are visible. This correlates with the statement of Barton & Wright (2000) that no upper limits of angles for avalanches seem to exist and is characteristic for Scotland. In comparison to other mountain ranges, no very high altitudes (highest points is on Ben Nevis with 1344 masl) and only little extremely steep slopes exist.

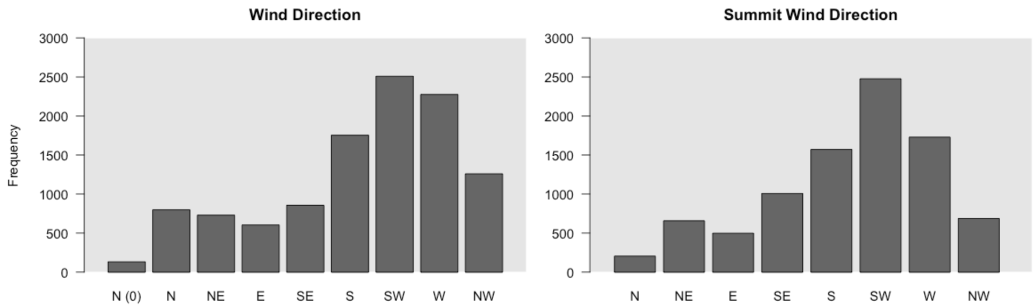


Figure 8.1: Distribution of wind directions measured at the snow profile locations and at the summit stations.

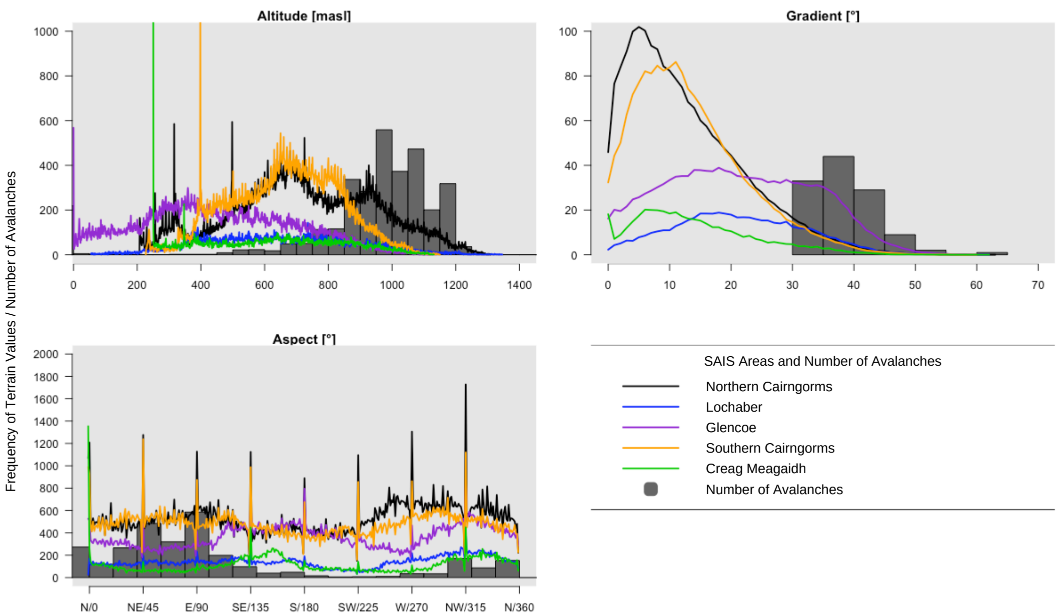


Figure 8.2: Comparison of available terrain in the five SAIS areas to the avalanche occurrences.

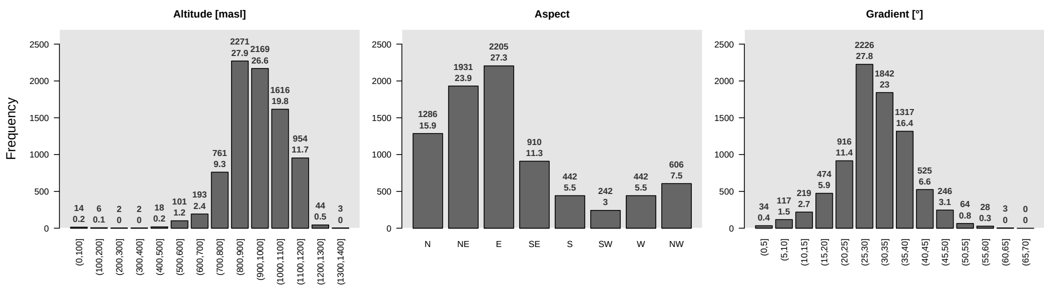


Figure 8.3: Terrain characteristics of snow profile locations.

Additionally, locations for high hazards can be described using the terrain information of the snow profile data (see figure 8.3), as they usually represent the most hazardous conditions of the day (Diggins 2010). General distributions look quite similar, but the peaks in the altitude and gradient are little lower lying at 800 masl to 900 masl and 25° to 30° in comparison to 900 masl to 1100 masl and 35° to 40° in the avalanche data set (see figure 7.4). This not only confirms the most dangerous avalanche terrain, but also shows that the SAIS does a good job in choosing most hazardous conditions for their snow profile locations.

However, the confinement of the reported avalanches to certain terrain features does not explain the constrain to the SAIS areas, as locations with similar terrain characteristics exist outside of the SAIS areas as well. Already Davison & Davison (1987, 54) mention that "one way of achieving [...] would be to identify areas of similar altitudes and terrain in Scotland". Figure 8.4 shows such an approach as coloured areas indicate every location in which the three terrain characteristics altitude, gradient and aspect lie within the typical ranges for avalanche activity. The thresholds are chosen rather narrow and the areas show therefore only the minimum extent for avalanche prone terrain.

The number of avalanches reported outside the SAIS areas is with 58 (less than 2% of all avalanches) small. The very first of these avalanches was reported in 1995, all of the others after December 2009 with an average of nearly 15 avalanches per season. This increased spreading of the avalanche occurrences beyond the borders of the SAIS areas since the season 2009/10 is visible in figure 7.1 and may be due to the fact that "higher visitor numbers may lead to overcrowding" (Hanley *et al.* 2001, 49) within the SAIS areas, or that it is ever getting less difficult to reach remote locations. Barton & Wright (2000, 131) mention the following example: "road travel is easier and the opening of Kessock and Skye bridges have significantly affected patterns of mountain usage. Climbing fashions change too and the icefall climbing on the back of Liathach is now well known to climbers from south of the border. All of this will inevitably affect the accident rate in the areas which have effectively become less remote".

The avalanches reported outside of the SAIS areas show no exceptional characteristics. However, of those in the Northern Highlands occurred each in similar terrain with about 200 meters distance to ridges (see figure 8.5 for some examples). They could either have been seen from the road or been reported by people who were next to the avalanche locations. The visibility from the road is only seldom mentioned in the comments of the reported avalanches and is analysed using a viewshed (see figure 8.6). Of the 15 avalanches which are reported in the Northern Highlands, 5 could have clearly been seen from the road, with the others it remains unclear. The majority of the avalanches reported between the SAIS areas show large distances to roads and are most likely not seen from a road. Most of the avalanches which are reported clearly southwards of the SAIS areas are next to roads and therefore probably be seen from a road. With two exceptions, all avalanches which are reported to have been seen from the road in the comments, lie within the SAIS areas.

The second and third statements of Ward (1984a) become important when trying to explain the strong constrain of the avalanche to the SAIS areas and he states that the distribution of the avalanches is highly dependent on where people are going: "some slopes appear to avalanche more often than others because they are more easily observed or because they gain something of a reputation which tends to be self-promoting" (Ward 1984a, 95).

He mentions for example 'Coire an Lochain' and 'Coire an-t Sneachda' as frequently avalanching locations. They belong to the Northern Cairngorms and are very popular for climbing (Fyffe

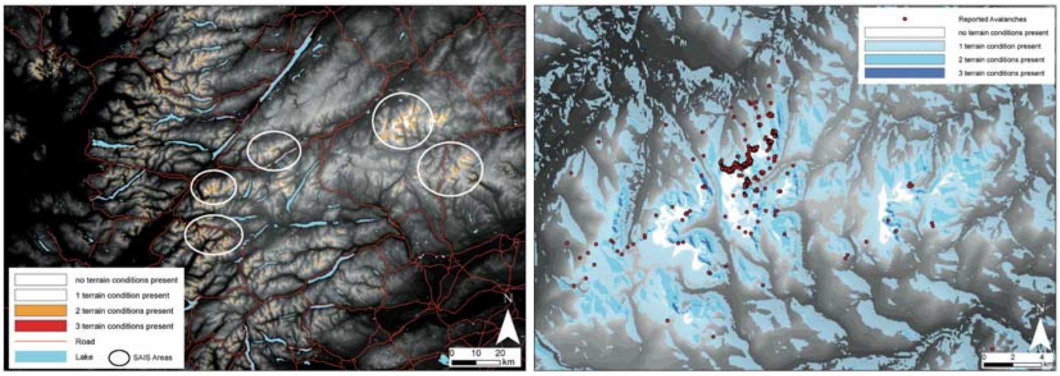


Figure 8.4: Distribution of the optimal terrain characteristics for avalanches (Thresholds: Altitude: 900 masl - 1100 masl, Aspect: N, NE and E, Gradient: 30° - 45°). Left: larger regions around SAIS areas | Right: Northern Cairngorms.

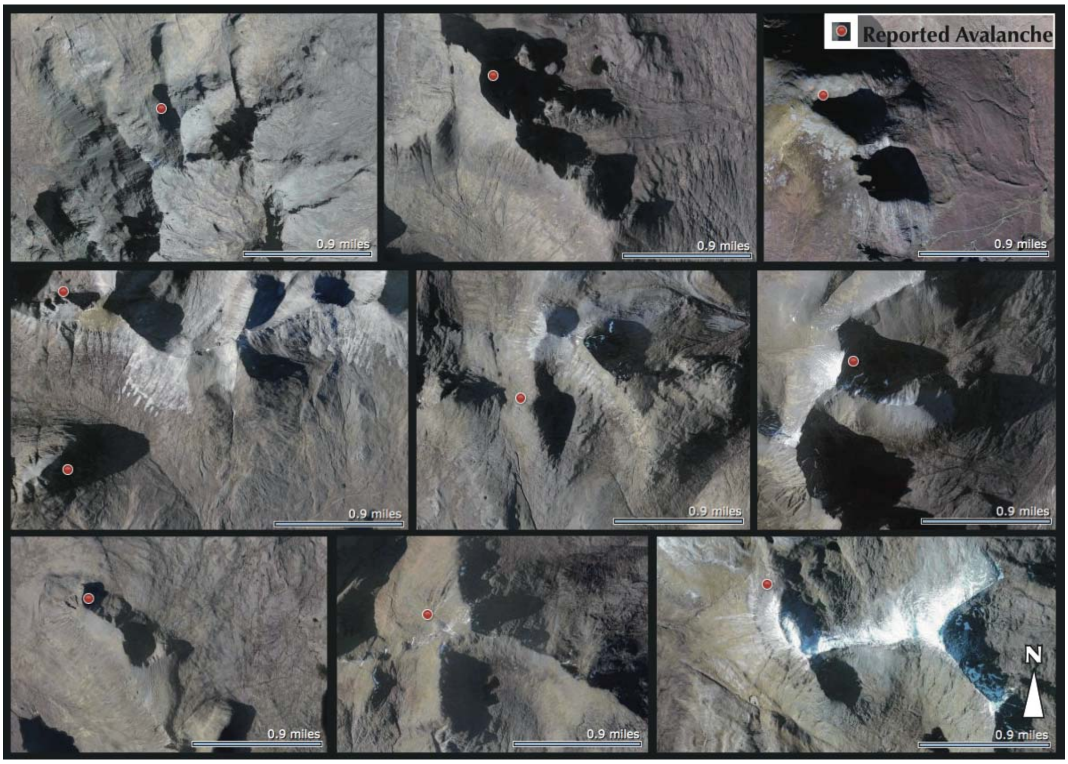


Figure 8.5: Examples of avalanches (red dots) reported in the Northern Highlands (SAIS 2013a).

2000). Figure 8.8 shows a comparison of the distribution of the avalanches with official climbing routes. The similarities of these two distributions confirm that "avalanche occurrences are recorded only where people can travel in the mountains or can see clearly from roads and paths" (Diggins 2010, 2). Human activity is not only related to mountaineers. "Most mapped avalanche occurrences are mainly located in the five SAIS operational areas, where SAIS forecasters are active" (Diggins 2010, 8). Additionally to the observers of the SAIS, other professionals such as the snow patrol or instructors are working in the Northern Cairngorms, Glencoe and Lochaber. After climbers, they are mentioned most frequently in the comments as having reported an avalanche (see figure 7.2). Further, it is even probable, that the reporting is done more seriously within the SAIS areas, as the forecasting service is better known by the public.

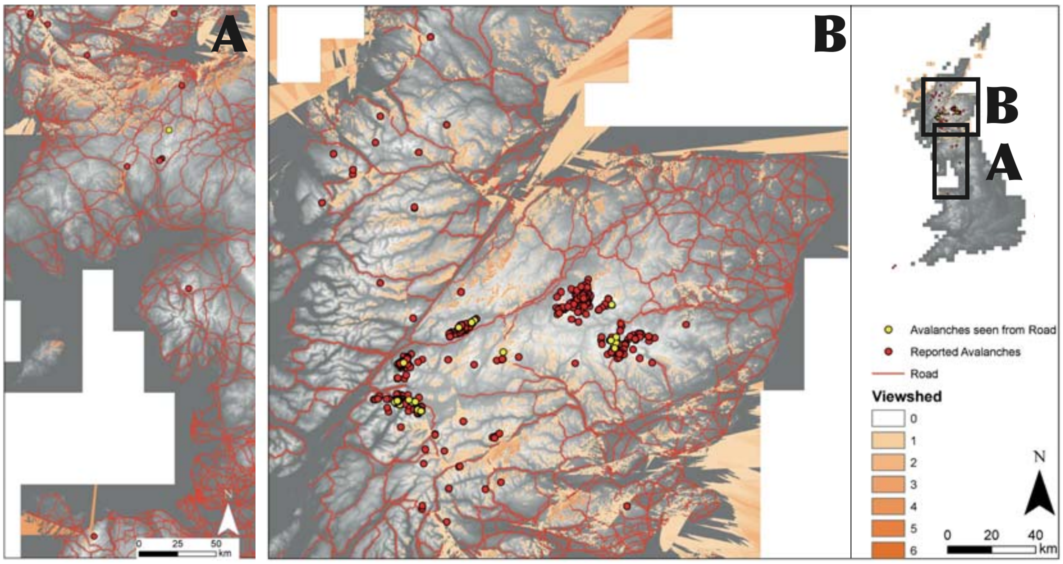


Figure 8.6: Viewshed of main avalanche areas.

These are the main bias present in the distribution of the reported avalanche data set and have always to be kept in mind. According to Diggins (2010, 2) “it can be assumed that a greater number of avalanche occurrences have taken place than have been recorded” and “it should not be concluded that the presented avalanche occurrences identify the main locations where avalanches take place”, as they “generally occur where human activity takes place” (Diggins 2010, 8).

Closely related to the human activity is the trigger (see section 2.1.2.5) as avalanches involving people are in 90% of the cases triggered by their victims (Diggins 2011b, Schweizer & Lütschg 2000, Schweizer & Lütschg 2001). For Scottish avalanches as well, this emphasises the importance of the human factor, but it is difficult to extract reliable information. The only attribute in the avalanche data set which includes information about the trigger type is ’Release’ and only for about 25% of avalanche occurrences, information is included. Only 111 (7%) avalanches are classified to be human triggered. When including the comment information, the number can be risen to almost 300, but also this corresponds to only 9% of all reported avalanches.

Nevertheless, some characteristics of human triggered, natural or cornice triggered avalanches are shown in figure 8.7. No avalanches are triggered by humans in November, and only very few natural releases take place in April. A higher proportion of avalanches is triggered by humans than has naturally released when drifting conditions (82% versus 68%) and no rain at 900 masl (95% versus 73%) are present. During the night, no avalanches are triggered by humans, but 80% between 9.30 am and 15.30 pm are human triggered in comparison to only 50% naturally in this timespan. Cornice triggered avalanches show similar values than naturally released avalanches.

Further information concerning humans, such as the number of people involved in an avalanche accident, the activity that has been carried out or the outcome of accidents are not included systematically in the data and are only mentioned in some comments. Comparisons to the results on the influence of the group size as presented in McCammon (2004) (see section 2.1.2.4) can therefore not be drawn. The only additional data sources that include similar information are Sharp & Whalley (2010) and the annual reports of the SAIS (Diggins 2010, Diggins 2011b, Diggins 2012). Figure 8.1 shows the numbers mentioned in the different data sources. The ranges are comparable but exact values differ in the majority of cases. Numbers are sometimes larger in the report, sometimes in the comments, and compared to Sharp & Whalley (2010), only less than 50% of the fatalities and less than 25% of the injuries are mentioned in the comments. Therefore the information retrieved from the comments hint towards some human aspects but exact numbers are not reliable. When assuming that the information included in the comments is correct, the values can be used as absolute minimum numbers. In reality, the amounts are probably much higher.

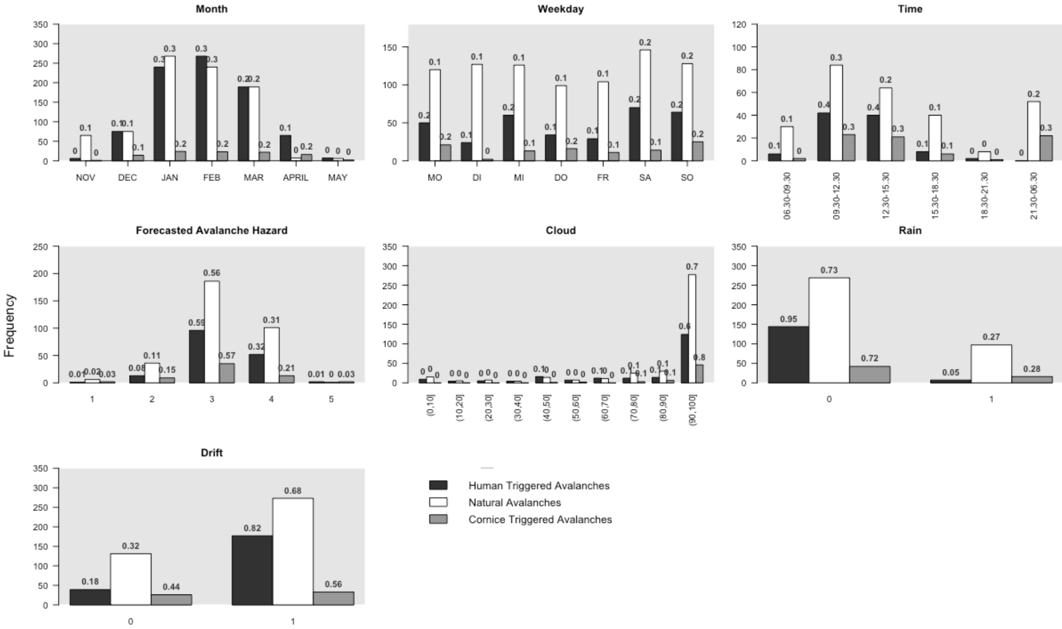


Figure 8.7: Comparison of natural released, human and cornice triggered avalanches for different attributes.

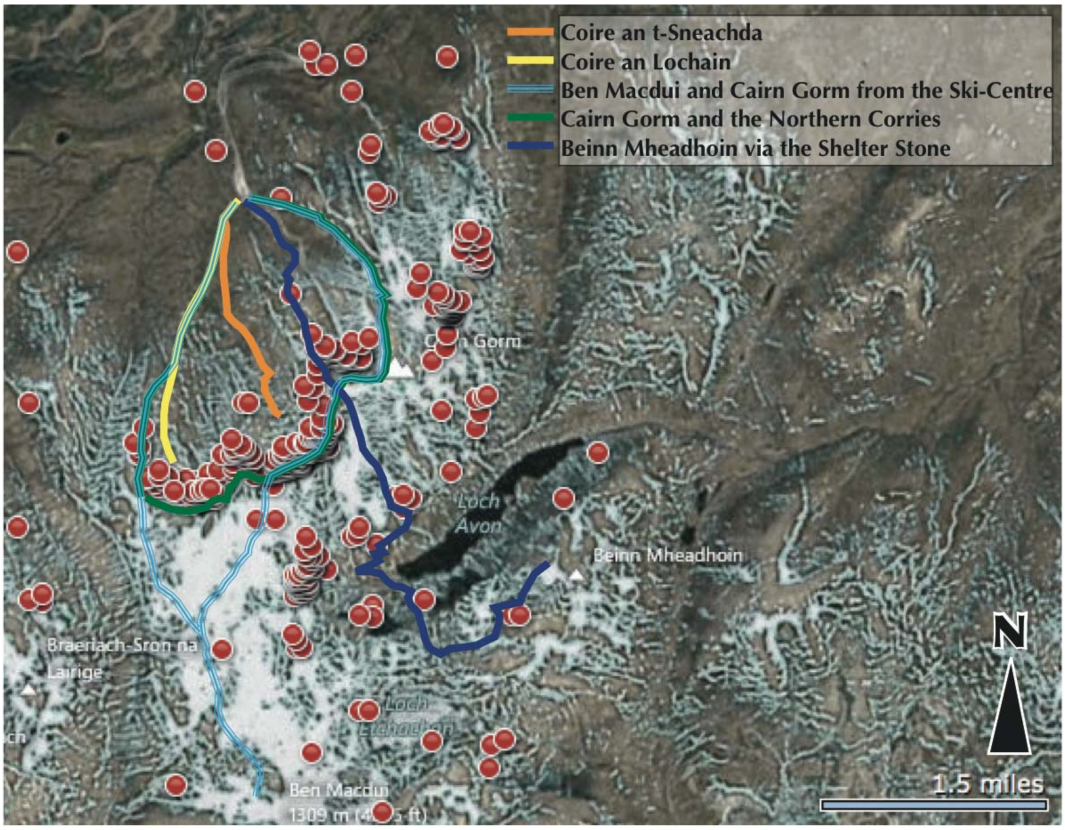


Figure 8.8: Comparison of official climbing routes and the distribution of avalanches in the Northern Cairngorms (routes: walkhighlands 2013a, background: SAIS 2013b).

Table 8.1: Comparison of data concerning human aspects from the annual reports of the SAIS (Diggins 2010, Diggins 2011b, Diggins 2012), Sharp & Whalley (2010) and from the comments in the reported avalanche data.

Source ---	Total Nr of Avalanches		Fatalities, Injuries and Burials			Trigger		Number of People Involved	
Season |	Report	Comments	Report	Comments	Sharp & Whalley (2010)	Report	Comments	Report	Comments
09/10	220	261	5	fatalities: 5 injuries: 4 burials: 2		176 natural	167 natural 51 human 14 cornice	44 persons 9: ski patrol / observers 14: skier / boarder 21: climbers / walkers	39 persons (+5*party) 8: ski patrol / observers 18: skier / boarder 81: climbers / walkers
10/11	178	179	1	fatalities: 0 injuries: 4 burials: 4		127 natural 11 cornice	117 natural 40 human 14 cornice	40 persons 7: ski patrol / observers 9: skier / boarder 24: climbers / walkers	29 persons (+4*party) 7: ski patrol / observers 11: skier / boarder 24: climbers / walkers
11/12	154	155	0	fatalities: 0 injuries: 2 burials: 6		127 natural	87 natural 33 human 37 cornice	33 persons 4: ski patrol / observers 6: skier / boarder 23: climbers / walkers	40 persons 7: ski patrol / observers 6: skier / boarder 23: climbers / walkers
1980-2009				fatalities: 24 injuries: 56	fatalities: 59 injuries: 231				

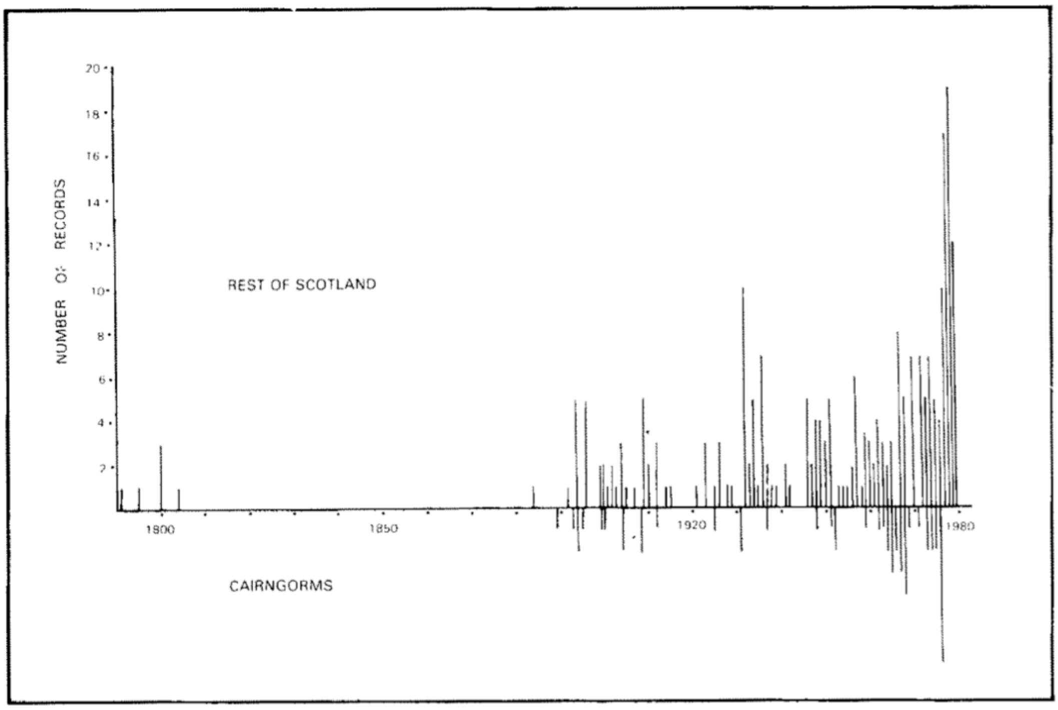


Figure 8.9: Number of avalanches recorded in Scotland between 1790 and 1980 (Ward 1984a, 101).

Temporal Distribution and Popularity

For the yearly distribution, it is not possible to make a direct comparison with Ward (1984a) because temporal coverages do not overlap with the current analysis. Ward (1984a, 101) found that since 1960 an increase in avalanche activity has taken place but "there is no reason to suppose that avalanches have suddenly become more common" because the number of avalanches strongly depends on observation and reporting. Ward (1984a) also states that peaks exist in the data, the biggest one in the year 1979 when 20 avalanches have been recorded. Comparing these numbers with the current results shown in figure 7.2, they are considerably smaller, as current average numbers per season are bigger than 100 and maximum frequencies reach even more than 200 avalanches per season.

The spatial spreading and multiple increased numbers are indications that not only the reporting of avalanches has become much more common, but that also the actual amount of avalanches has increased. As winter sports are gaining popularity, more people are out on the Scottish hills each winter, raising automatically the avalanche activity. With increasing popularity, the accessibility of the mountain ranges improves and the mountaineers have the possibility to use infrastructures in place. This enables winter sporting with less efforts and a broader range of people is heading towards the mountains. Already in ski areas, where tickets have to be bought, it is difficult to estimate average numbers of people per season, as a high variability is present (Holden 1998). Determining the amount of recreationists out in the Highlands is even much more difficult. However, according to Diggins (2010), the numbers of visits of the SAIS webpage can give a clue about the amount of people that are visiting the area (given in table 8.2). Each of the four last seasons shows at least 200'000 visits per winter, reflecting the absolute smallest number of recreationists.

Table 8.2: Numbers viewing the daily SAIS avalanche forecast reports (Diggins 2010, Diggins 2011b, Diggins 2012).

AREA	2011/12 SEASON	2010/11 SEASON	2009/10 SEASON	2008/9 SEASON
Northern Cairngorms	51,902	89,236	111,573	66,813
Southern Cairngorms	28,045	44,936	69,492	37,855
Lochaber	46,615	63,653	92,188	65,848
Creag Meagaidh	25,976	40,066	63,744	37,843
Glencoe	38,279	59,073	95,907	67,937
Mobile	25,836			
TOTALS	216,653	296,964	432,904	276,296

Ward (1984a, 103) summaries that the season can last for 9 months (September to April) and that "peak months appear to be February, March and April, with December and January less important". The current avalanche data shows a clear peak of avalanche activity in January, February and March with April and December being less important and only few avalanches occurring in November and May. The trend mentioned by Ward (1984a) that the habits in winter climbing are changing and people do no longer wait for the milder spring months to go out, probably continues. Nevertheless, the average season length of about 150 days stays similar, which means that on average at least 1300 people per day are out in the backcountry. However, a considerable number of days per season is characterised by bad weather and "a disproportionate share of avalanche accidents occur during blue-sky days" (Fredston *et al.* 1994, 476). Further, a tendency that more people are going out on weekends exists. In the avalanche occurrences per day (see figure 8.10), especially the number of Saturdays, but also Sundays and Wednesdays show higher numbers in comparison to the rest of the week, getting more pronounced the more avalanches release per day. The weather and irregular distribution of the number of recreationists over weekdays can raise the daily number of people who are in the hills many times. Also HIE (1996) in Purves *et al.* (2003, 344) and Hanley *et al.* (2001, 36) speak of larger numbers with an estimated 767'000 mountaineers who visited the Scottish Highlands and Islands in 1996.

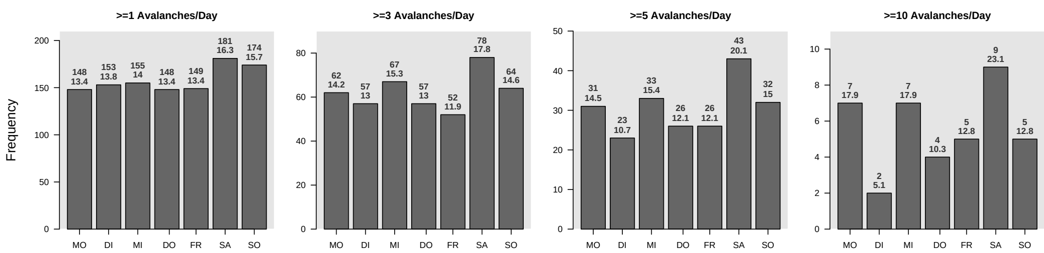


Figure 8.10: Distribution of avalanche days on weekdays considering different thresholds for an avalanche day.

Considering the time of the day at which avalanches release, in the current data set only little information is available and Ward (1984a) had the same problem. Out of the 22 avalanches he analysed "fell 17 in the afternoon, one at noon, two in the morning and two overnight" (Ward 1984a, 103). He mentions that the data should be used carefully especially because these times could also "reflect the time at which most climbers reach the climbing area" (Ward 1984a, 103). In the current data, the time can additionally describe the moment at which the avalanche has been reported instead of the actual moment of release. The time is available for about 85% of all avalanches and it is remarkable that nearly 17% of these release during the night. This can probably either be explained by the uploading time or by typing errors.

Avalanche Types, Magnitudes and Importance of the SAIS

Until 1950, thaw avalanches in spring time were the only avalanche type thought to be of interest. However, "windslab avalanches of the snowier months, January and February, have been regarded as a chief danger" some years later (Ward 1984a, 103). These two avalanche types make up an important part of the total avalanche amount (between 50% and 70%, see table 8.3). The current pattern analysis shows additionally the importance of avalanches which release during cold conditions and in relation to weak snow covers, as about 20% to 30% of the total amount of reported avalanches occurs during these situations.

An other distinction is made between dry and wet avalanches (see figure 8.11). Wet avalanches more frequently release in the early afternoon, when it is raining at an altitude of 900 masl or when no drifting takes place. In contrast, about 90% of the dry avalanches release when it is not raining at 900 masl or when drifting conditions are present. Wet avalanches release naturally in about 90% of the cases, dry avalanches are triggered by humans in 30% of the cases.

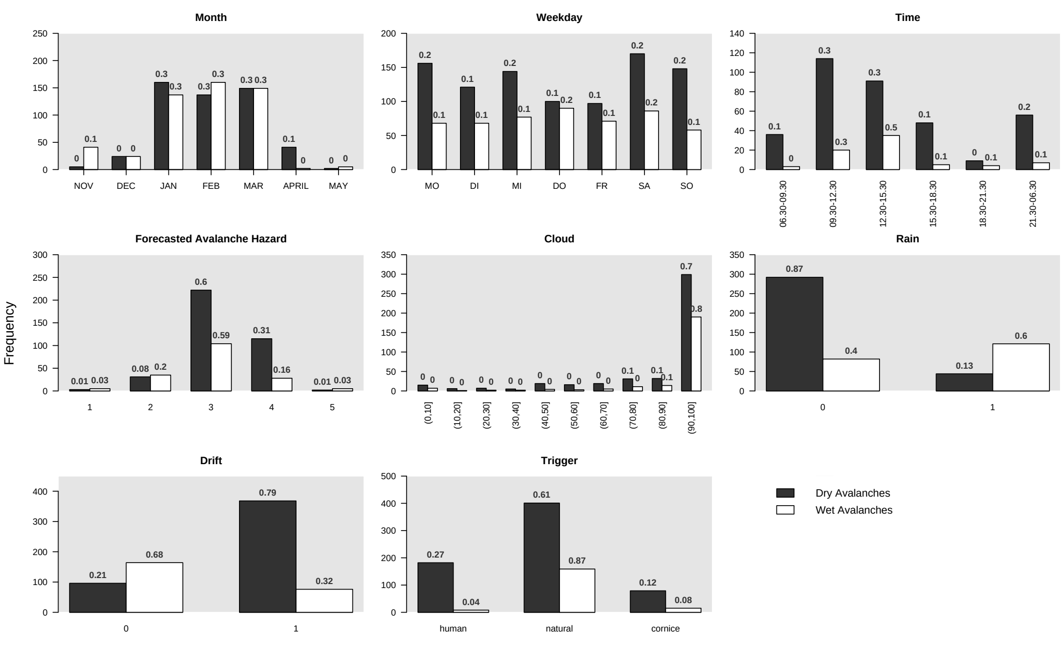


Figure 8.11: Comparison of dry and wet avalanches considering other attributes.

About other characteristics of avalanche types (see section 2.1.1) no information is included in the data and it cannot be differentiated between full-depth and surface avalanches, which is already a main concern around 1990 in the research of Davison & Davison (1987, 51): "in Scotland, only three full-depth avalanches have been recorded definitely, so our knowledge of this type of avalanche is very limited."

The distribution of the avalanche size confirms the conclusion of Ward (1984a, 100) "that although the majority of Scottish avalanches are undoubtedly small, exceptions occur which produce very large events".

The maximum number reached in the current data for the fall length corresponds to two historical cases from 1979 in 'Glen Geusachan' and 1959 in 'Coire na Ciste' which Ward (1984a) presents in his paper. However, it cannot be concluded without any doubt whether the high numbers of more than one kilometre are typing errors or rare but realistic exceptions of particularly large avalanches.

According to Ward (1984a) fracture depths are mostly between 0.5 and 1 meter, they can also reach three or four meters and even one avalanche has been reported with a fracture depth of 8 meters (Ward 1984a). The highest fracture depth in the current data set lies at 3 meters, and identical to the findings of Ward, the majority of avalanches show fracture depths smaller than one meter.

Average widths of the avalanches found by Ward (1984a) are about 50 meters, but he adds that much broader ones are also possible. The current analysis brought similar results with nearly 90% of avalanches showing maximal widths of less than 50 meters and about 12% with widths between 50 and 100 meters. Five avalanches are reported to be wider than 500 meters (500, 600, 700, 750 and 1000 meters).

Summarising, even though the majority of avalanches in Scotland is relatively small, the reported avalanches have shown that exceptionally large ones can occur occasionally as well. In relation to the spatial spreading of avalanche releases, it can also be possible that in future even more large events have to be expected. A general increase in avalanche hazard is also by Harrison *et al.* (2001) proposed, who analysed changes in snowfall patterns in Scotland in relation to climate change. Conditions during the last years have become less reliable and "heavy storms and rapid thaws had increased the avalanche risk" (Harrison *et al.* 2005, 3). The longer the more, infrastructure could become an issue in Scotland as well. In February 2010, one of the first times a major rail line was blocked due to an avalanche and it was "probably the first time explosives have been used to safeguard infrastructures from avalanche threat in the UK" (Diggins 2010, 7). Such a development would mean a completely new problem for Scotland, as, until recently, avalanches were 'only' of concern for recreation. Until today, the aim of the SAIS consisted mainly of working against the 'Scarcity' heuristic mentioned by McCammon (2002), so that their forecasts are not seen as prohibitive but as useful and necessary support for mountaineers. Further and different approaches than already used by the SAIS would become important. This means, that the work of the SAIS is not finished yet and will not finish tomorrow. Further effort is needed and the work of the SAIS is gaining importance.

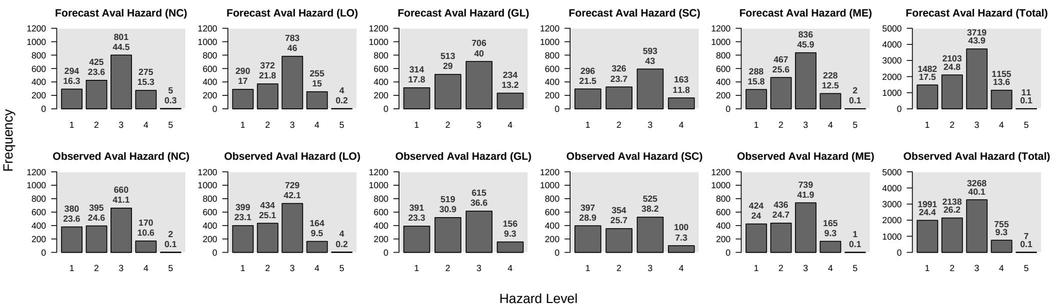


Figure 8.12: Frequencies of forecasted and observed avalanche hazard levels for each of the five SAIS areas and in total.

Already today, the importance of the SAIS becomes obvious, and figures 8.12 and 8.13 confirm the statements of Diggins (2011b, 3) that "the percentage of hazard levels shows a reasonable consistency between all five areas and demonstrates that for a significant part of the winter the hazard level is *Considerable* or *High*". Hazard level 5 (*Very High*) has been used only in the areas Northern Cairngorms, Lochaber and Creag Meagaidh, was forecasted a total of 11 times and observed a total of 7 times. Hazard level 3 (*Considerable*) is in all areas most frequently used. The forecasted avalanche hazard levels are generally higher than the observed ones which means that the SAIS is rather on the safe side and tends to overestimate the hazards (in 24% of cases). In 70% of cases, the forecasting is correct and in only 6%, the forecasted hazard level is lower than the observed one.

A key component which has to be considered when assessing the properties of the avalanche data set, is the form on the SAIS webpage which can be used to report an avalanche (see figure 8.14).

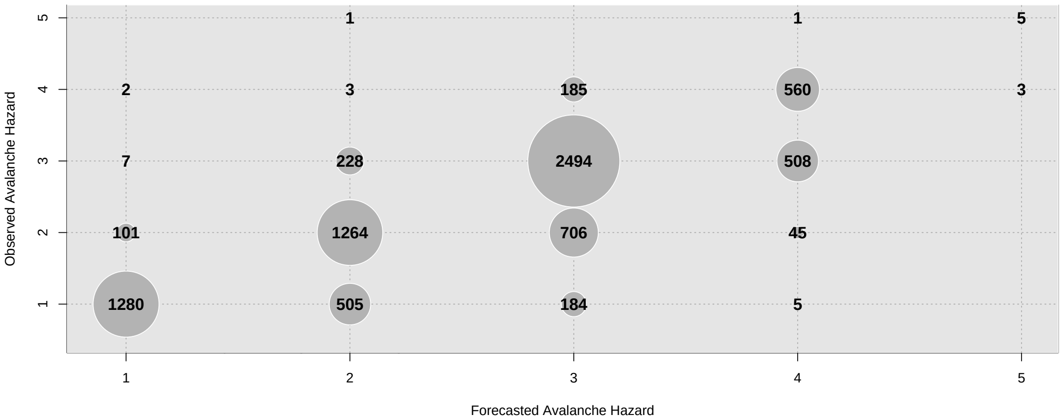


Figure 8.13: Forecasted versus observed avalanche hazard levels.

Originally, a large form had to be filled out including information about the type of the avalanche, meteorological conditions and information about victims (Version 1). Since 2007, the form is much shorter (Version 2), and its characteristics seem to influence the distribution of certain avalanche attributes.

Many values have to be chosen from a drop down list leading to an obvious pattern in the data with large peaks at round numbers (see figure 8.15 for two examples). The red bars show extremely frequent numbers such as *50*, *100*, *150*, *200*, *300*, *400* and *500* for the fall length or *1*, *10*, *20*, *30*, *40*, *50*, *100* for the maximum fracture depth. Blue bars are less extreme but peak also at round numbers. The majority of values therefore are rough estimates that give an indication of ranges, but exact numbers are not reliable. This is not surprising, considering the source of this data set and the circumstances in which the data originates. Scottish avalanches are not measured with a ruler, their dimensions are estimated either directly in the field, or much more likely hours later, when sitting at home in front of the computer filling out the form and trying to remember the situation by heart.

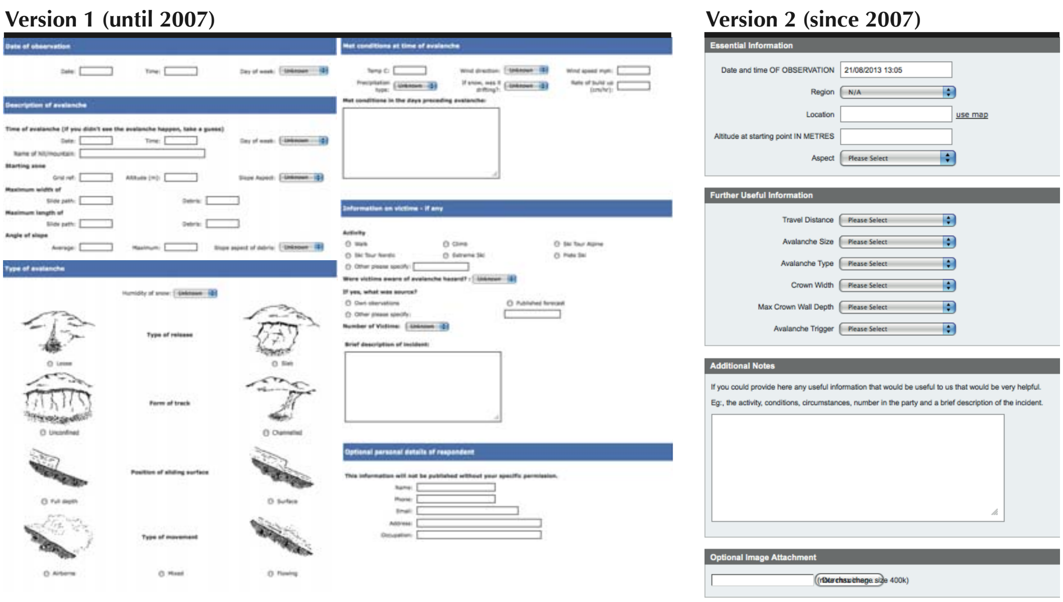


Figure 8.14: Two versions of the form on the webpage of the SAIS for reporting an avalanche (SAIS 2013b: Report an Avalanche).

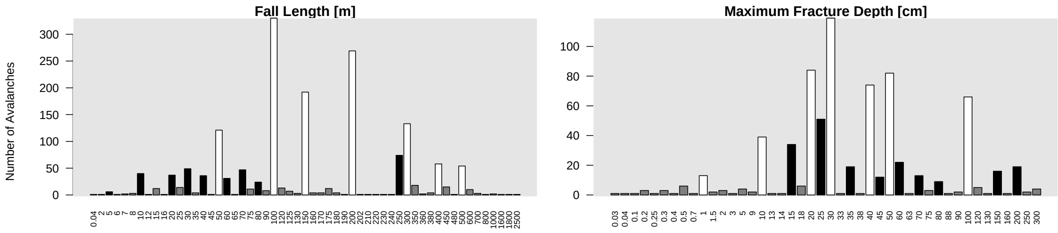


Figure 8.15: Original distribution of the attributes 'Fall Length' and 'Maximum Fracture Depth'.

Summarising ...
the findings of R. Ward which are 30 years old and were done without any major reliable data source, could well be confirmed with the current analysis. New characteristics have emerged and further insights are provided. General knowledge about avalanche characteristics and theories, which have perhaps subjectively been known by the SAIS, could be verified and examined in greater detail as well, using an objective analysis based on broad data. Therefore the main aim of the first part of this thesis could be achieved. For sure, the statement of Ward (1980, 31) that "the victims are either mountaineers or cross-country skiers as the problem is confined to the steeper and remoter slopes" becomes relative. These steeper and remoter slopes seem to increase the longer the more in number and actually become less remote because the behaviour of recreationists, the material and the infrastructure develops. A tendency in the data emerges, that especially the spatial distribution of avalanche occurrences relevant to the SAIS will increase in the future and that winter activities are no longer confined to the popular sites located within the SAIS areas.

These considerations lead to the following answers to the research questions:

What can be said about general avalanche activity in Scotland?

The general avalanche activity in Scotland is widespread. Thus, the avalanche problematic plays an important part in the Scottish winter and the SAIS plays a major role since 1988. In comparison to the findings of Ward (1984a), the spatial distribution of the avalanches seems to spread and the frequency of reported avalanche occurrences has increased during the last thirty years. The spatial clustering of the avalanches in this data set is very strongly related to human activity and it has to be kept in mind that the used data set includes not the entire amount of avalanches that has released, but only those avalanches which have been reported.

Which characteristics do the Scottish avalanches show?

The distribution of the avalanches is on the one hand spatially confined to the terrain, mainly altitude and gradient, with maximum frequencies at about 1000 masl and between 35° to 45°, and on the other hand temporally restricted to the winter season which lasts from November to May with clearly the most avalanche occurrences between January and March. Several important avalanche types exist, including not only thaw and windslab avalanches as has been stated by Ward (1984a), but also cold avalanches that are related to weak snow cover. The majority of avalanches is small, with fracture depths of less than one meter, widths of less than 50 meters and fall lengths below 100 meters in 50% of the cases. However, exceptions of avalanches with significantly larger dimensions can occur as well.

8.2 Part B: Combination of Data Sets

The explorative method applied to the available data sets motivated by KDD using a variety of visualisations and iterative processing steps (see section 6.1) has proven to be a good choice. A broad overview of the data could be derived in part A of this thesis, it was possible to cope with the nature of the data and a high working quality could be ensured as no blackbox algorithms had to be used and every single working step could be tested. The snow profile data could be assessed detailed, its limitations recognised early and overcome in this part of the thesis by combining the different data sets into a new one which provides a reliable data basis for part C of this thesis and further studies. The main focus in this part lies on the data quality and by assessing the four main dimensions of the data quality (see section 6.2), the data could be analysed comprehensively.

Which properties do the different data sets show?

Each data set shows different characteristics, includes a single combination of attributes and time coverages. There is no data set available which covers the total amount of information.

Which limitations does the data have?

The data shows considerable amounts of missing information, typing errors in different attributes occur, and consistency within as well as between attributes is not always ensured.

Is it possible to combine the different data sets into one comprehensive data set?

Yes, the combination of the different data sources has been successful. Nearly each attribute has been combined from more than one data source. The selection of the data source could sensibly be done in any case.

Is it possible to ensure satisfying data quality in the resulting data set?

The two resulting data sets show a satisfying quality which could be increased considerably in comparison to the original data sets. Typing errors could be removed, consistency within and between attributes could be ensured and time coverages for each attribute are as complete as possible.
However, some limitations concerning completeness are still present. Especially in the beginning of the operation of the SAIS between 1990 and 1994 and between the years 2003 to 2007 missing values are present. The ones in the early beginning are not that significant as they are missing only occasionally. The ones between 2003 and 2007 are due to reorganisations within the SAIS. In some files stored in combination with the ProfileObs data on the FloyerDVD, information about paper records for these years is included:

- NC: for 2003 to 2007 paper records in the SAIS data repository exist.
- LO: no records between 2003 and 2007 exist.
- GL: no records between 2003 and 2006, for 2006-2007 paper records in the SAIS data repository exist.
- SC: for 2003 to 2007 paper records in the SAIS data repository exist (for 2005-2006 only March and April).
- ME: for 2006-2007 paper records in the SAIS data repository exist.

Therefore for some areas and some timespans it should be possible to further minimise the amount of missing data.
Nevertheless, the two resulting data sets can now be used for future projects without the necessity of any further basic data preparation. Therefore the main aim of this second part, including the analysis of the available data and the combination into one comprehensive data set, could be achieved.

8.3 Part C: Pattern Analysis

Resulting Patterns, Their Importance and Reliability
The pattern analysis is based on various scenarios, different input data sets are considered and a validation including several aspects is done. Parts of the results are obvious and reliable, others have to be interpreted carefully as only little information is available. The following four situations are included in the majority of scenarios and therefore resulted as very clear and obvious patterns:

- **Thawing situations:** mean summit air temperatures are above 0 °C and strong short-time warming of in average +3 °C to +6 °C takes place. Clearly the least weaknesses exist in the snow cover. With rising air temperatures, snow temperatures increase as well, the snow gets wet, probably influenced by liquid water and destabilises.
- **Cold periods:** mean summit air temperatures stay below -5 °C and a Weakness Index with an average of 7 to 8 is present in the snow cover. Weak layers develop preferably when persistent cold temperatures exist and can destabilise the snow cover considerably.
- **Drifting conditions with much new snow:** very high wind speeds with means of 60 mph and up to 300 mph for the seven day sum as well as high amounts of new snow (6 to 8 and up to 50 in the seven day sum) are present. This leads to strong drifting conditions and the accumulation of dangerous windslabs.
- **Not extreme conditions:** low summit air temperatures, considerable winds, some new snow and weaknesses tend to be present, but no extreme values are reached by any of the variables. The different factors can be more or less pronounced, depending on the input data.

It is difficult to extract very detailed descriptions from the available data and many uncertainties are still included, especially in interpretations related to the information from the comments. Nevertheless, it is possible to distinguish between the three very obvious patterns including thaws, cold periods and situations with strong drifting and new snow.
Table 8.3 shows an overview of the total amount of avalanches that are assigned to each cluster for the scenarios with >=1 avalanches/day for each SAIS area and in total. More than 95% of the total amount of avalanches belonging to one of the five SAIS areas can be assigned to one of the above mentioned patterns, except for the Northern Cairngorms where the amount lies little lower at 91.5%. Avalanches which could not be assigned to a cluster are either due to missing snow profile information at the avalanche day (see appendix A.6.3) or to a too large amount of missing data in the snow profile information for the Northern Cairngorms, Creag Meagaidh and the total data (see figure 7.15). Between 5% and 30% of the avalanches are assigned to the cluster which shows not extreme conditions. Therefore about 70% to 95% of the avalanches can be related to one of the first three obvious patterns. They seem to be most important considering Scottish avalanche activity and can be seen as main hazardous situations. A similar amount of 60 out of 73 avalanches (~82%) could be related to typical avalanche weather from Ward (1984b) as well.

Table 8.3: Number of avalanches that are assigned to each cluster for the >=1 avalanches/day scenarios (T: Thawing pattern, C: Cold pattern, D: Drifting pattern, N: Not extreme pattern, (): not strongly pronounced.

Area	Cluster 1			Cluster 2			Cluster 3			Cluster 4			Cluster 5			Cluster 6			Total Avalanches Assigned to Clusters	Total Avalanches in Area	% of Total Avalanches Assigned to Clusters
	Pattern	[abs]	[%]	Pattern	[abs]	[%]	Pattern	[abs]	[%]	Pattern	[abs]	[%]	Pattern	[abs]	[%]	Pattern	[abs]	[%]			
NC	T	110	**16.7**	C	49	**7.4**	C	144	**21.9**	D	228	**34.6**	N	128	**19.4**	-	-	-	659	720	91.5
LO	T	188	**19.4**	C	195	**20.1**	D	321	**33.1**	(D)	247	**25.4**	-	-	-	-	-	-	971	976	99.5
GL	(T)	173	**37.1**	N	50	**10.7**	D	132	**28.3**	(D)	111	**23.8**	-	-	-	-	-	-	466	469	99.3
SC	T	49	**26.9**	C	43	**23.6**	(D)	23	**12.6**	D	54	**29.7**	N	13	**7.1**	-	-	-	182	189	96.3
ME	T	240	**35.2**	C	42	**6.2**	C	177	**26.0**	D	64	**9.4**	(D)	110	**16.2**	N	48	**7.0**	681	689	98.8
TOTAL	T	543	**21.2**	C	235	**9.2**	D	618	**24.2**	D	345	**13.5**	N	818	**32.0**	-	-	-	2559	3043	84.1

The CA used to establish these patterns is a widely used technique that allows to comprise complex data into a manageable amount of information (Burns & Burns 2008, Pang-Ning *et al.* 2006).

Disadvantages of the technique include that results can always be reached, even if the method is not appropriate or not applied correctly. A high amount of the analysis and especially the interpretation is subjective and decisions that are made on input data or algorithms can influence the results heavily (Backhaus *et al.* 2008, Mooi & Sarstedt 2011). Therefore it is closely looked at these negative aspects and the influences of the choices of the threshold for an avalanche day, order of input data, handling of outliers and missing values as well as the applied algorithm are analysed and different input data sets for each SAIS area and the total amount of data are considered by comparing results (see section 7.3.2.5). Summarising, differences are relatively small in any case and the clusters do not change their general characteristics. The analysis is based on broad data and represents a time period of more than 20 years. Thus, the resulting patterns seem to be stable and it can be said with good confidence that the patterns are a reliable representation of the available data. They summarise the broad variety of data in such a way that major properties are described and weight is laid on the most important aspects. The main aim of the third part of this thesis, which includes the searching for patterns in the available data set, could therefore be achieved.

Scottish Patterns Compared to Other Ones

The results can be compared to other studies (see section 2.3.2). Referring to Scotland, R. Ward is the only author who described situations which can be seen as patterns. In Ward (1980), he mentions especially *'thaws'* and *'cold periods after heavy snow falls'* as main hazardous situations, in Ward (1984b) *'fresh snowfall', 'cold spells'* and *'thaws'* are summarised as major dangerous conditions, and in Ward *et al.* (1985) *'cold periods after heavy snowfall', 'thaws'* and *'storms'* are differentiated (see section 3.5.1). The cold spells and thaws can clearly be confirmed with the current analysis, which is based on a different methodical approach and on much more data than has been available about thirty years ago. Especially the crucial role of cold periods and the development of weak layers consisting for example of depth hoar which was seen as being "unlikely to be a major cause of avalanches in Scotland" by Ward (1980, 36) could be emphasised. Storms cannot clearly be distinguished as it is not easy to define thresholds. However, the third pattern describing the drifting situation shows exactly the two main features which characterise a storm: high winds and snow falls. A further important point which has not been considered by R. Ward in any of his publications, is the fact that probably also a pattern has to be considered, which is characterised by not extreme conditions. Depending on the area, a considerable amount of the avalanches (between 5% and 30%) is assigned to this pattern, still introducing uncertainties in the avalanche hazard.

A cluster 'Catastrophe' as mentioned in Harvey *et al.* (2002) and Signorell (2001) is not obviously present in the current results. However, similar to storms, catastrophes can emerge if the drifting pattern is very pronounced and a weak snow cover exists at the same time. Mainly two factors seem to contribute to the absence of a distinct 'Catastrophe' pattern in the current analysis. First, in the end of the applied CA it is mainly looked at the cluster centres and depending on cluster sizes, extreme values of outliers which could represent 'Catastrophe' situations become relativised. Second, the bias of the reported avalanche data set may play an important role as less people are out in the hills in stormy conditions leading to less avalanches being seen and reported (Moss 2009). However, this also means that avalanches during storms are more likely to be of natural release and less important for the SAIS.

The other patterns found in the CA of Harvey *et al.* (2002) and Signorell (2001) are similar to the current ones, however, variables are slightly different assigned. The *'rise in temperature'* in the current thawing pattern is strongly correlated to generally warm temperatures. *'Unstable snow cover, [...], low temperature'* represents exactly the current cold pattern and the *'strong winds'* cluster is in Scotland in the majority of cases related to heavy snowfall and therefore a combination of Harvey's clusters 2 and 4. A pattern *'high amount of new snow and low temperature'* without strong winds is in the current analysis not pronounced and was only in the Northern Cairngorms distinguished.

Similarities can be seen when comparing the current results to the ten published patterns in Tirol as well (see section 2.3.2). The results of this thesis are a good summary, even though the current analysis is not done as detailed and the patterns are more general. *'Cold following warm / warm following cold'* and the *'spring time scenario'* are partly included in the current thawing pattern. The current analysis did only show *'warm following cold'*, as strong cooling could not be distinguished and the *'spring time scenario'* is not constrained to the spring months in Scotland, but can occur during the whole winter. The two patterns *'surface hoar blanketed with snow'* and *'graupel blanketed with snow'* describe weak layers and are included in the current cold pattern, but they cannot not be distinguished as detailed. *'Cold, loose new fallen snow and wind'* is clearly represented by the current drifting pattern, but temperatures compared to the Alps can be higher in Scotland. *'Snowfall after a long period of cold'* is partly included in the current cold pattern but cannot not be described as detailed and also patterns such as *'the second snowfall'*, *'snow-poor zones in snow-rich winters'* or *'sliding snow'* go beyond the scope of this thesis.

The patterns proposed by the SLF are more similar to the current ones. They are more general as those of Tirol, but are primarily based on the characteristics of the snow without taking into account specific weather factors. Nevertheless, the *'wet snow pattern'* clearly represents the current thawing pattern and the *'drifted snow pattern'* the current drifting pattern. The *'old snow pattern'* can best be correlated to the current cold pattern as weak layers play an important role in both. The *'new snow pattern'* is not as clear distinguished in the current results, as new snow in Scotland is very often related to high wind speeds and therefore included in the drifting pattern.

Some input variables such as the total snow depth, snow temperature or snow water equivalent often included by other authors in similar approaches (see e.g. Birkeland *et al.* 2001, Jomelli *et al.* 2007, McCollister *et al.* 2003 and Stoffel *et al.* 1998 in section 2.3.2) are not considered in the current CA because of the available data. However, the snow temperature is in combination with further important factors (see section 2.2.2.2) considered in the Weakness Index.

Summarising

The current analysis is very well comparable with already established similar studies and some clear differences which are related to the special Scottish conditions can be seen. High amounts of new snow are more strongly related to high winds which leads to a larger influence of drifting processes, and thaws are not a typical spring situation but can occur during the entire Scottish winter.

The pattern analysis was motivated by the SAIS, as they would like to publish patterns on their webpage (pers.comm.: Mark Diggins, Co-ordinator of the SAIS). For the first time, an approach in this direction could be realised and the basis for such a publication can be provided with this thesis. The three main hazardous situations are found with high certainty and can be used as simple, well understandable and relatively easy recognisable typical hazardous situations during a Scottish winter.

For a publication, an even more detailed description of the single patterns would be helpful as it might be possible that the cold pattern, for example, can be divided into 'cold after new snowfall' and 'cold for more than 7 days'. Especially important would a further analysis of the characteristics of the 'not extreme' patterns be as they still introduces a high amount of uncertainty in the avalanche problematic. Situations with not extreme conditions are most difficult to recognise and include for up to 30% of all avalanches large uncertainties. When no obvious hazardous pattern is recognisable, the heuristics mentioned by McCammon (2002) and McCammon (2004) become especially important and the 'Familiarity' as well as the 'Consistency' heuristic can influence the behaviour of recreationists heavily, leading preferably to human triggered avalanches and accidents.

However, it is possible that the avalanche days in this pattern are characterised by factors which are not included in the current analysis. More complex relationships between variables or iteration processes can take place which were not able to be extracted by the used CA. For example, freeze-thaw cycles mentioned as stabilising factor in the annual reports of the SAIS could not be considered and detailed temporal sequences such as snowfall after long cold periods (compare Ward 1980 or Ward *et al.* 1985 in section 3.5.1) or storms in spring time (compare annual SAIS reports in section 3.5.4) neither.

Implications

Therefore some further considerations have to be done. Eleven variables which in fact represent only three different factors (air temperature, wind speed, amount of new snow and stability of the snow cover) are included in the analysis. Further variables such as total snow depth or rain are not considered. Seven out of the eleven input variables relate on measurements taken at the summit stations. The summit attributes are included by the SAIS into the data set since they are thought to be representative for the whole area (Diggins 2010). Nevertheless, the summit measures do not exactly represent the conditions at the avalanche locations and therefore influence the results. The only other possibility would be to deduce the data from the snow profiles to the avalanche locations, as in the reported avalanche data set no information about weather conditions is included. Exemplarily tested in this thesis is a simple regression using a temperature gradient mentioned in Barton & Wright (2000) of 1 °C / 200 m for the Northern Cairngorms. The air temperatures measured at the snow profiles are calculated to the height of the summit station and compared with the summit station measurements (see figure 8.16). No better relation is present than using the air temperatures measured directly at the snow profile location.

It has already been mentioned by Buchan (1890, XXii) that temperature gradients are highly variable: "rates observed from day to day differ widely from the means and from each other" and "temperature observations are affected by very local site conditions" (McClatchey 1996, 37). The temperature gradient mentioned by Barry (1992) in McClatchey (1996, 39) is with -6 °C to -10 °C even higher and he states that "temperature lapse rates are steep in Britain due to the maritime air mass characteristics". Due to these uncertainties, an interpolation has not been applied in the current analysis.

Considering wind speeds, it might even be much more complex, as they show higher temporal and spatial variability and are strongly dependent on the local terrain. Price (2000, 163) mentions that "dimension, relief, slope and aspect all play their role in affecting the flow of air across a mountainous terrain" or that "the shape of the mountain, and particularly the summit, is likely to influence the recorded wind-speed characteristics, as will the exact position of the anemometer on the summit" (Price 2000, 164). Also Roy (1983, 3) states that "the winds at the summit were very strongly affected by the topography of the mountain". Additionally, wind speeds show a very high variability between years (Dawson 2009) and also large altitudinal differences can be present: "the average wind speeds at the summit AWS are nearly twice those observed at even very exposed low level stations" (McClatchey 1996, 43).

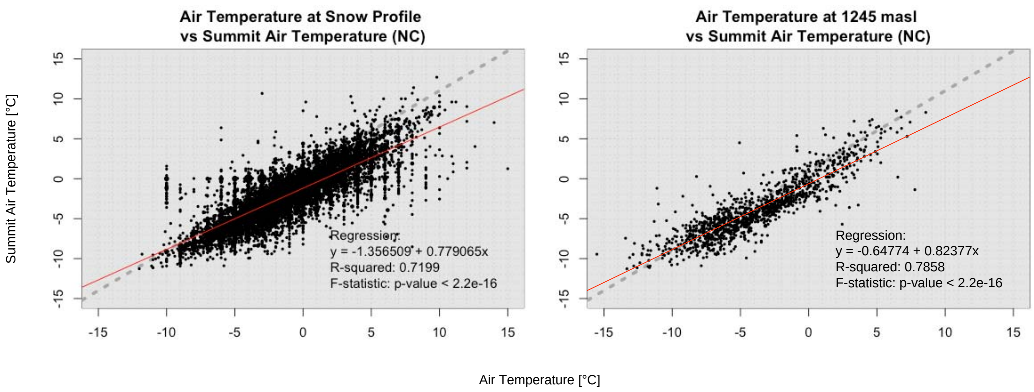


Figure 8.16: Comparison of the attributes 'Air Temperature' and 'Summit Air Temperature' for the Northern Cairngorms.

But on the other hand, the measurements at the summit stations are averages over longer time periods, compared to the measurements taken at the snow profile locations which only represent one specific time of the day. Therefore the summit measures are probably more representative for the whole day and are more reliable as they include several measurements. However, this implies, that summit measures are mean values, calculated from maximum and minimum air temperatures or wind speeds. Using only these mean values in the analysis introduces simplifications. Processes in warm situations are, for example, mostly dependent on maximum air temperatures, but during prolonged cold periods, minimum air temperatures can play a decisive role as well.

Therefore the use of summit measures as done in this thesis is currently and without having much additional effort the most reliable choice and has led to reliable results, in spite of the simplifications and loss in detailed information that are introduced.

Additionally, more complex patterns such as freeze-thaw cycles have not been possible to be included and also situations such as inversions, which are playing an important role during Scottish winters (Buchan 1890, McClatchey 1996), could not be considered. Time series and the history before an avalanche day are with the three and seven day sums only partly included.

However, with additional interpretations, it is possible to extract some of these aspects. It is exemplarily looked at the number of days with negative air temperatures that preceded an avalanche day in the cold clusters of the optimal scenarios summarised in tables 7.17 to 7.22. In the cold cluster of Lochaber, 12 out of 17 avalanche days lie within a period of constantly negative air temperatures. The number of days with negative air temperatures preceding the avalanche days range between 4 and 40 and show an average of about 17 days. For Creag Meagaidh 14 out of 18, for the Northern Cairngorms 16 out of 18 and for the Southern Cairngorms 9 out of 10 avalanche days lie within a period of negative air temperatures with corresponding means of 9.7 days, 16.5 days and 8 days that have preceded the avalanche days. Therefore the majority of the avalanche days belonging to the cold pattern are indeed within longer periods of cold temperatures. In the thawing cluster of Lochaber, 19 out of 35 avalanche days are first days showing positive air temperatures after several days with negative air temperatures (in average 9 days). Therefore at least one part of the thawing patter include the typical situation *'cold conditions followed by strong thaws'* mentioned in the annual reports of the SAIS.

The Weakness Index seems to have only little influence on the results, but the number of weaknesses is always minimised in warm situations and maximised in very cold situations. The Weakness Index and also the air temperature differences make up only 2/11 of the input variables, the three others (summit air temperature, summit wind speed and snow index) are represented each with three different measures. Therefore the influence of the Weakness Index and the air temperature difference is smaller and could with a new design of the analysis be improved. For example, an additional weighting of the input variables could be included in the CA (Backhaus *et al.* 2008, Pang-Ning *et al.* 2006). The Weakness Index itself is a strong simplification of the processes as well. Especially influences such as the spatial variability or effects on different scales (Schweizer *et al.* 2008) could not be considered. Nevertheless, referring to literature mentioned in section 2.2.2.2, the Weakness Index is based on an approach often used and seems to summarise the stability of the snow cover well.

These aspects lead to the following answers to the research questions:

Which patterns does the data show?

The data shows three very clear patterns including a thawing situation with summit air temperatures above 0 °C and strong short time warming, cold periods with mean air temperatures below -5 °C and drifting conditions, in which high wind speeds and large amounts of new snow are present. A further pattern includes not such extreme conditions and seems to be the most difficult for predicting as hazards can arise also in non-obvious situations when complex interactions between several factors are taking place.

Are there typical factors or combinations of factors concerning topography, weather, snow cover and humans which lead to an increased avalanche activity?

The patterns are characterised by certain factors and combinations of factors that highly influence the avalanche activity and increase the avalanche hazard. For non-avalanche days, similar factors play a role, but in a contrasting way. In summary, the analysis has shown that the contributing factors cannot be limited to one or two and that no single factor can be used to draw conclusions about avalanche activity. Several different factors and especially their combinations are determining the avalanche hazard which has already been concluded by Harvey *et al.* (2002).
Concerning topography, certain ranges of altitude (mainly between 900 masl and 1100 masl) and gradient (mainly between 30° and 60°) constraining avalanche activity can be defined. Aspect is not that limiting, but most avalanches clearly release on northern slopes. However, under certain conditions and especially in relation to weak snow covers, exceptions occur as well.
Most important weather factors include the amount of new snow, the air temperature and wind speed as highest avalanche activities correlate with extremely high or low temperatures, high amounts of new snow and high wind speeds. In the snow cover, especially weak layers play an important role, mostly consisting of unstable crystals such as facets or depth hoar. Last but not least, also the human factor has proved to be important in reporting as well as triggering avalanches. Especially in Scotland, where the avalanche problematic is currently constrained to recreation, humans cannot be ignored. Nevertheless, their influence is difficult to estimate and the human factor stays probably the most decisive but also the most unpredictable one.

9 CONCLUSIONS

Avalanches are an issue in Scotland and make up an important part of the Scottish winter; this could be shown in this thesis. A variety of different data sets have been analysed in a broad methodical approach, an overview of avalanche activity is provided, a structured data basis is established and patterns describing most hazardous conditions are being found. Research on avalanches is world-wide important. However, research on Scottish avalanches has only limited been done. Since the overview of the avalanche activity from R. Ward in 1984, not many further studies have been published. Nevertheless, many people have been concerned with the Scottish avalanche hazard. The SAIS was established and over a nearly 30 year period, a considerable amount of data has been collected. Therefore the data basis, background knowledge, methods as well as the behaviour of recreationists have developed considerably during the last years.

This thesis is a new step in the research history of Scottish avalanches, provides insights into different aspects of the state of the current avalanche problematic and supports the SAIS hopefully for future work. In each of the three parts, several aims are achieved and some major insights are gained.

9.1 Achievements

In part A:

- An overview of the current state of avalanche activity is provided.
- This overview is as current as possible and is based on data covering a time period from the beginning of the operation of the SAIS to the end of the last season, May 2013. The five most popular winter sport areas in Scotland, 'Northern Cairngorms', 'Lochaber', 'Glencoe', 'Southern Cairngorms' and 'Creag Meagaidh', are taken into account. Descriptions of Scottish avalanches and their major characteristics are provided. Various aspects of the spatial and temporal distribution, avalanche magnitudes and terrain characteristics of avalanching slopes are described. Although the influence of the human is difficult to extract from the data, related aspects are analysed and the importance of the human factor is emphasised.
- The findings of R. Ward are mostly confirmed, additional aspects are being considered and new insights provided. Knowledge and theories on which the work of the SAIS is based, have previously often been subjectively known by the observers and can be confirmed and enhanced in this thesis, using an objective analysis based on broad data.

In part B:

- The unstructured and chaotic data base, consisting of various data sets with different characteristics, is combined into one comprehensive data set.
- This new data set shows a satisfying data quality. It is as complete and consistent as possible. Typing errors and inconsistencies are removed and the time span covered by each attribute is maximised.
- The data set is described in detail with metadata and made available for the SAIS.
- The time consuming, tedious and sometimes annoying data pre-processing steps are done very detailed and carefully for the whole data that is available from the beginning of the SAIS operation until the end of the season 2012/13. Many problematic issues which accumulated through the last twenty years, are resolved and removed. In future, it will be much easier to hold the data quality steadily at the current state, as such an intensive pre-processing will not be necessary any more.
- Additional material including further results, tables and plots of the data analysis is stored in a systematic way on a DVD which is available for the SAIS. This information can be used for further studies.

In part C:

- A basis for the publication of Scottish avalanche patterns is established.
- Days showing high avalanche activity are grouped into clusters that describe most hazardous conditions, using a new approach never applied to the data before.
- Variables including information about air temperature, wind speed and amount of new snow as well as the stability of the snow cover are considered. The building of sums over several days allows to consider the temporal history preceding an avalanche day to some extent.
- The data is analysed for each SAIS area separately, to account for spatial differences, and for the total amount of data.
- Several intensities of the avalanche activity as well as days on which no avalanches occurred are considered.
- Most obvious hazardous patterns are found, including thawing situations, long cold periods and drifting conditions with large amounts of new snow. Depending on the input data, between 70% and 95% of the avalanches can be assigned to one of these patterns. Additionally, 5% to 30% of the avalanches belong to a pattern showing no extreme conditions.
- Differences to areas in other countries such as Tirol or Switzerland are shown by comparing the current results to already established patterns. Most important differences include that storms are not clearly distinguished as a single pattern, that new snow is more strongly related to heavy snowfalls and that thaws can occur during the whole winter in Scotland.
- Patterns or most hazardous situations mentioned by R. Ward are updated, confirmed and enhanced. Especially the importance of cold periods is emphasised.
- The resulting patterns can be used as basis for developing a set of patterns which can be published by the SAIS to support recreationists in assessing the avalanche hazard in various situations.

9.2 Insights

Additionally to the achievements mentioned above, some more general insights can be provided.

- The thesis shows that avalanches are a crucial part of the Scottish winter. The importance of the work of the SAIS during the last thirty years is emphasised as the publication of forecasts is an essential service to many recreationists. Also in future, avalanches will be an issue. Developments will take place and the SAIS will have to adapt. The analysis has shown tendencies in an increase not only in the number of reported avalanches but also in their spatial extent. It could be possible that the avalanche problematic will broaden and that the longer the more also infrastructures and questions of artificial releases and preventive measures will become important.

- The importance of the human factor concerning avalanche activity becomes obvious again, but also the difficulties in estimating the influences get apparent.
- The findings of R. Ward from thirty years ago can generally be confirmed. Only few statements have to be contradicted, the knowledge is enhanced and further aspects are considered. Additionally, some of Ward's results which were based on a very small and not entirely reliably data base, are confirmed using broader data and different approaches. Therefore his work is a very useful basis and can be seen as valuable reference also for further projects, showing the state of the knowledge of thirty years ago.
- Concerning the pattern analysis, it got apparent again that not only one single factor determines the avalanche hazard, but that the combination of different factors is crucial. Relationships and interactions of the weather, snow cover, terrain and human are complex and in every study can only a small part of the whole processes be described.
- The cluster with not extreme conditions has shown that still many uncertainties exist and that the processes related to avalanche activity are of high complexity. It is not possible to classify every avalanche in a systematic way in some few typical situations showing clear characteristics, as a rest risk in not typical conditions is always present and has to be considered by recreationists.
- A new step in the history of Scottish avalanche research is reached and for the first time, the broad variety of different data sets is used and combined. Additionally, the thesis is a contribution to the international research on avalanches. Results are similar to previous studies that have been conducted in other countries using comparable approaches. Many of their conclusions can be confirmed and new insights are provided.
- The application of a broad methodical and explorative approach motivated by KDD has proved to be appropriate. The characteristics of the data would have not allowed only statistical analysis. With the combination of GIS, various visualisations and statistics, different perspectives on the data are provided. Similarily concluded McCollister *et al.* (2002, 116) as well: "there is much potential for the concepts of KDD and GVis to improve utilisation of historic weather and avalanche databases".

9.3 Recommendations for the SAIS

The data analysis of the avalanche data set and the snow profile information has shown some aspects which could be further improved during the data collection or when post-processing the data. Most of the below mentioned points simplify primarily the basic data preparation steps if the data will be used for future projects.

- The frequent data collection is important and should continue. The snow profiles combined with stability tests in the field and information from automatic weather stations provide broad information.
- The form on the webpage for reporting avalanches influences the characteristics of the data. Drop down lists are a good service to the reporter and can help in estimating certain attributes, but the range of possible answers is constrained. Especially when being unsure, one tends to choose randomly instead of recording the lack of knowledge. It might be valuable to include an option 'unsure' in the drop down lists.
- If the focus on human triggered avalanches should be enlarged in the future, it would make sense to systematically collect related data. The easiest method would be to include relevant details in the avalanche reporting form; at least the number of people who were involved and the activity that has been carried out to hold the form short and simple.
- It might improve the amount of available data within the attribute 'Gradient', when the gradient in a similar way than altitude or aspect is included in the current reporting form.
- It could be helpful to record the results of stability tests not within the comments, where also further aspects are described, but as additional attribute. The results of Rutschblock and Shovel Shear tests occur in a considerable amount of the snow profiles and make up the biggest

content of the comments. The effort to extract this information could be much simplified when introducing two new attributes.

- Unified shortenings in the comments in snow profile data could help to simplify further analysis. For example for 'easy shear quality' resulting from the shovel shear test are the following abbreviations included in the current data sets: *Easy S/T, Easy ST, EST, Easy Sh, easy shear, easy shovel, easy shovel shear, easyST, easy SS, E/S, e shear, easy shear 2, S Shear 2, SHEAR 2, S/shear 2, s/s 2, shovel shear 2, shovel/S 2, SST 2*. This makes an extraction of the information very time consuming and introduces uncertainties.
- An objective measure including the amount of new snow, given for example in centimetres, could allow better comparisons with other studies and thresholds such as the critical new snow depth could be applied to the data.
- It might be interesting and useful to collect data about the snow water equivalent or density of new snow as this variable is used in several other studies and is there seen as an important factor.
- The data is up until the end of the season 2012/13 structured and typing errors and inconsistencies are as far as possible removed. To hold this data quality in future at the current state as well, it could make sense to invest after each season some hours post-processing the new data. Typing errors and inconsistencies can be removed and the data can be stored in a consistent way without a big effort. This will prevent an accumulation of limiting data characteristics over several years.

9.4 Further Research

This thesis provides a broad overview of various aspects of the current avalanche activity in Scotland. Most important characteristics of the available data are assessed, but there are still factors which would be interesting to be analysed in greater detail or to be analysed additionally.

Considering the overview of the avalanche activity, it has not been looked at the spatial distribution of single attributes. Exact locations of the snow profiles are considered neighter, as they are seen as being representative for the total SAIS areas (Diggins 2010). However, it might be interesting for the SAIS to compare the exact snow profile locations to the avalanche locations to further improve the choices on where to take a snow profile.

The cluster analysis has proven to be a good choice to start a pattern analysis and three very obvious hazardous situations are found with high certainty. However, only the aspects included in tables 7.17 to 7.22 are analysed for the resulting patterns and a very detailed description is not provided. It would be interesting to further analyse the characteristics of the days assigned to the patterns, especially for the ones where not extreme conditions are present.

It is possible that these avalanche days indeed released during conditions which are not as extreme as in the other patterns. However, it is likely that factors which could not be considered in the current analysis are playing an important role as well. A close look at the distribution of the attributes of these days could give first insights or single days could exemplarily be analysed, as for example done in Signorell (2001). Further, already available methods which have previously been applied to the Scottish avalanche data to support the forecasts might be used to analyse the characteristics of these avalanche days. These include SIRE, an Information Retrieval System (Purves & Sanderson 1998), Cornice (Purves *et al.* 2003, Purves *et al.* 1998b) or Support Vector Machines (Posdnoukhov *et al.* 2008 and Posdnoukhov *et al.* 2011). A further possibility could include a similar approach as used by Eckert *et al.* (2010) who analysed an avalanche cycle in December 2008 in the French Alps which was characterised by a severe storm including large amounts of new snow, considerable drifting and strong temperature variations, using a numerical snow cover model to analyse abnormal situations.

Additionally, the chosen CA is not entirely flexible and could be improved further. The issues mentioned in section 8.3*: Implications* could tried to be resolved. The method could be developed in a way that further attributes such as the binary variables 'Rain' or 'Drift' can be included, the variables could be weighted (e.g. Backhaus *et al.* 2008) or an other approach instead of the Weakness Index could be used. The measurements taken at the summit stations which are primarily used in the current analysis could tried to be related to the locations of the avalanches. A simple regression has not been a valuable improvement, but further methods such as geographically weighted regressions (see e.g. Fotheringham & Charlton 1998 or Lloyd 2010) in combination with expert knowledge might provide better results. To assign each avalanche day to one cluster is a simplification and the characteristics of the avalanche days are reduced to those of the clusters. Therefore, it could be valuable to test a fuzzy clustering approach in order to assign avalanche days only with a certain probability to the clusters (for example proposed by Steinley 2006).

To overcome the still missing data in the resulting data set of part B, information of the paper records between 2003 and 2007 could be included. Further missing values in various attributes which are not systematically missing could tried to be estimated, for example, with expert knowledge. Other studies show successful applications (e.g. Eckert *et al.* 2010). This could enable the further analysis mentioned above or might be a first step for the application of a PCA on the data which could not be done as initially intended in this thesis. The PCA could be used, similar as in other successful studies, to assign situations characterised by high avalanche activity to synoptic atmospheric patterns (Birkeland *et al.* 2001, Esteban *et al.* 2005b, García *et al.* 2008) and to analyse anomaly fields (Casteller *et al.* 2011, Mock & Kay 1992, Rangachary & Bandyopadhyay 1987). With such an approach, avalanche days could be related not to patterns as have resulted in this thesis, but to typical Scottish atmospheric patterns.

BIBLIOGRAPHY

Adams, L. (2005). *A Systems Approach to Human Factors and Expert Decision-Making within Canadian Avalanche Phenomena*. Ph.D. Thesis, Royal Roads University.

Albers, S., Klapper, D., Konradt, U., Walter, A. & Wolf, J., eds. (2006). *Methodik der empirischen Forschung*. Springer DE.

Amaratunga, D., Baldry, D., Sarshar, M. & Newton, R. (2002). Quantitative and Qualitative Research in the Built Environment: Application of “Mixed” Research Approach. *Work study*, 51 (1), 17–31.

Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2008). *Multivariate Analysemethoden - Eine anwendungsorientierte Einführung*. Springer.

Ballantyne, C. (1989). Avalanche Impact Landforms on Ben Nevis, Scotland. *The Scottish Geographical Magazine*, 105 (1), 38–42.

Barraclough, T., Blackford, J. & Zaiser, M. (2010). Snow in Scotland: Snowmicropen Analysis of Natural and Artificial Snow Samples. *In: International Snow Science Workshop, Squaw Valley, California, USA.*, 367–372.

Barry, R. (1992). *Mountain Weather and Climate*. Routledge, London.

Barton, B. & Wright, B. (2000). *A Chance in a Million? Scottish Avalanches*. Scottish Mountaineering Trust.

Barton, J. & Borthwick, A. (1982). August Weather on a Mountain Summit. *Weather*, 37 (8), 228–240.

Birkeland, K., Mock, C. & Shinker, J. (2001). Avalanche Extremes and Atmospheric Circulation Patterns. *Annals of Glaciology*, 32 (1), 135–140.

Bryden, D., Westbrook, S., Bums, B., Taylor, W. & Anderson, S. (2010). Assessing the Economic Impacts of Nature Based Tourism in Scotland. *Report*, Scottish Natural Heritage Commissioned Report No. 398.

Buchan, A. (1890). *The Meteorology of Ben Nevis*. Robert Grand & Son, Edinburgh and Williams & Norgate, London. Transactions of the Royal Society of Edinburgh.

Burns, R. & Burns, R. (2008). Cluster Analysis. *In: Business Research Methods and Statistics using SPSS*, Sage, Chapter 23.

Burrough, P. (2001). GIS and Geostatistics: Essential Partners for Spatial Analysis. *Environmental and Ecological Statistics*, 8, 361–377.

Burrough, P. & McDonnell, R. (1998). *Principles of Geographical Information Systems*. Oxford University Press Inc., New York.

cairn gorm (2013). Cairn Gorm Data. *http://cairngormweather.eps.hw.ac.uk/ [Accessed: 26.06.2013].*

cairngorm mountain (2013). Climbing. *http://www.cairngormmountain.org/activities/climbing [Accessed: 31.03.2013].*

Casteller, A., Villalba, R., Araneo, D. & Stöckli, V. (2011). Reconstructing Temporal Patterns of Snow Avalanches at Lago del Desierto, Southern Patagonian Andes. *Cold Regions Science and Technology*, 67, 68–78.

Davison, R. & Davison, S. (1987). Characteristics of Two Full-Depth Slab Avalanches on Meall Uaine, Glen Shee, Scotland. *Journal of Glaciology*, 33 (113), 51–54.

Dawson, A. (2009). *So Foul and Fair a Day: A History of Scotland's Weather and Climate*. Birlinn, Edinburgh.

Diggins, M. (2010). Report for Winter 2009 / 10. *Report*, Sportsscotland Avalanche Information Service.

Diggins, M. (2011a). Interpreting Snow Profiles. *www.sais.gov.uk [Accessed: 20.06.2013].*

Diggins, M. (2011b). Report for Winter 2010 / 11. *Report*, Sportsscotland Avalanche Information Service.

Diggins, M. (2012). Report for Winter 2011 / 12. *Report*, Sportsscotland Avalanche Information Service.

Eckert, N., Coléou, C., Castebrunet, H., Deschatres, M., Giraud, G. & Gaume, J. (2010). Cross-Comparison of Meteorological and Avalanche Data for Characterising Avalanche Cycles: The Example of December 2008 in the Eastern Part of the French Alps. *Cold Regions Science and Technology*, 64, 119–136.

Eckert, N., Parent, E., Belanger, L. & Garcia, S. (2007). Hierarchical Bayesian Modelling for Spatial Analysis of the Number of Avalanche Occurrences at the Scale of the Township. *Cold Regions Science and Technology*, 50 (1), 97–112.

esri (2013). ArcGIS. *http://esri-germany.de/products/arcgis/index.html [Accessed: 14.06.2013].*

Esteban, P., Jones, P., Martin-Vide, J. & Mases, M. (2005a). Application of Multivariate Statistical Techniques for the Characterisation of the Surface Circulation Patterns Related to Heavy Snowfalls in Andorra, Pyrenees. *International Journal of Climatology*, 25, 319–329.

Esteban, P., Jones, P., Martin-Vide, J. & Mases, M. (2005b). Atmospheric Circulation Patterns Related to Heavy Snowfall Days in Andorra, Pyrenees. *International Journal of Climatology*, 25, 319–329.

Fayyad, U., Piatetsky-Shapiro, G. & Smyth, P. (1996). The KDD Process for Extracting Useful Knowledge from Volumes of Data. *Communication of the ACM*, 39 (11), 27–34.

Ferguson, R. (1985). High Densities, Water Equivalents, and Melt Rates of Snow in the Cairngorm Mountains, Scotland. *Weather*, 40 (9), 272–277.

Fierz, C., Armstrong, R., Durand, Y., Etchevers, P., Greene, E., McClung, D., Nishimura, K., Satyawali, P. & Sokratov, S. (2009). *The International Classification for Seasonal Snow on the Ground*. IHP-VII Technical Documents in Hydrology N°83, IACS Contribution N°1, UNESCO-IHP, Paris.

Fischer, M. & Wang, J. (2011). *Spatial Data Analysis: Models, Methods and Techniques*. Springer, Berlin Heidelberg.

Floyer, J. (2010). Report: Scottish Snow and Avalanche Database. *Unpublished Report*.

Föhn, P. (1992). Characteristics of Weak Snow Layers or Interfaces. *In: International Snow and Science Workshop, Beckenridge, Colorado, USA*, 160–170.

Fotheringham, A.S., Brunsdon, C. & Charlton, M.E. (2000). *Quantitative Geography: Perspectives on Spatial Data Analysis*. SAGE Publications Ltd.

Fotheringham, A.S. & Charlton, M.E. (1994). GIS and Exploratory Spatial Data Analysis: An Overview of some Research Issues. *Geographical Systems*, 1, 315–327.

Fotheringham, A.S. & Charlton, M.E. (1998). Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis. *Environment and Planning A*, 30, 1905–1927.

Fredston, J., Fesler, D. & Tremper, B. (1994). The Human Factor-Lessons for Avalanche Education. *In: International Snow Science Workshop, Snowbird, Utah, USA*, 473–487.

Fyffe, A. (2000). *Winter Climbs in the Cairngorms*. Cicerone.

García, C., Marti, G., Oller, P., Moner, I., Gavalda, J., Martínez, P. & Pena, J. (2008). Atmospheric Patterns Leading Major Avalanche Episodes in the Eastern Pyrenees and Estimating Occurrence. *In: International Snow Science Workshop, Whistler, British Columbia, Canada*, 97–105.

geograph (2013). Geograph Coverage Map. *http://www.geograph.org.uk/mapbrowse.php [Accessed: 17.09.2013].*

glenmore lodge (2013). Glenmore Lodge - Scotland's National Outdoor Training Centre. *http://www.glenmorelodge.org.uk/ [Accessed: 30.04.2013].*

Good, I. (1983). The Philosophy of Exploratory Data Analysis. *Philosophy of Science*, 50, 283–295.

Goodchild, M., Haining, R. & Stephen, W. (1992). Integrating GIS and Spatial Data Analysis: Problems and Possibilities. *International Journal of Geographical Information Systems*, 6 (5), 407–423.

Gordon, J. (1993). The Cairngorms. *Quaternary of Scotland*, 6, 259–276.

Greene, E., Birkeland, K., Elder, K., Lazar, B., Moore, M., Sharaf, D., Tremper, B. & Williams, K. (2010a). Avalanche Observations. *In: Snow, Weather, and Avalanches: Observation Guidelines for Avalanche Programs in the United States*, The American Avalanche Association, Chapter 3, 69–83.

Greene, E., Birkeland, K., Elder, K., Lazar, B., Moore, M., Sharaf, D., Tremper, B. & Williams, K. (2010b). Manual Snow and Weather Observations. *In: Snow, Weather, and Avalanches: Observation Guidelines for Avalanche Programs in the United States*, The American Avalanche Association, Chapter 1, 3–20.

Greene, E., Birkeland, K., Elder, K., Lazar, B., Moore, M., Sharaf, D., Tremper, B. & Williams, K. (2010c). Snowpack Observations. *In: Snow, Weather, and Avalanches: Observation Guidelines for Avalanche Programs in the United States*, The American Avalanche Association, Chapter 2, 21–68.

Grung, B. & Manne, R. (1998). Missing Values in Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, 42, 125–139.

Guptill, S. & Morrison, J. (1995). *Elements of Spatial Data Quality*. Elsevier.

Guy, Z. (2011). *The Influence of Terrain Parameters on the Spatial Variability of Potential Avalanche Trigger Locations in Complex Avalanche Terrain*. Master's Thesis, Montana State University, Bozeman, Montana.

Guy, Z. & Birkeland, K. (2012). Avalanche Trigger Locations in Complex Terrain. *In: International Snow Science Workshop, Anchorage, Alaska*, 458–461.

Guy, Z. & Birkeland, K. (2013). Relating Complex Terrain to Potential Avalanche Trigger Locations. *Cold Regions Science and Technology*, 86, 1–13.

Hägeli, P. & McClung, D. (2003). Avalanche Characteristics of a Transitional Snow Climate - Columbia Mountains, British Columbia, Canada. *Cold Regions Science and Technology*, 37, 255–276.

Hägeli, P. & McClung, D. (2007). Expanding the Snow-Climate Classification with Avalanche-Relevant Information: Initial Description of Avalanche Winter Regimes for Southwestern Canada. *Journal of Glaciology*, 53 (181), 266–276.

Haining, R. (2003). *Spatial Data Analyis: Theory and Practice*. Cambridge University Press.

Hand, D. (1998). Data Mining: Statistics and More? *American Statistician*, 52, 112–118.

Hanley, N., Begona, A. & Shaw, W. (2002). Rationing an Open-Access Resource: Mountaineering in Scotland. *Land Use Policy*, 19 (2), 167–176.

Hanley, N., Koop, G., Begona, A., Wright, R. & Nevin, C. (2001). Go Climb a Mountain: An Application of Recreation Demand Modelling to Rock Climbing in Scotland. *Journal of Agricultural Economics*, 52 (1), 36–52.

Harrison, J., Winterbottom, S. & Johnson, R. (2001). Climate Change and Changing Snowfall Patterns in Scotland. *Report*, The Scottish Executive Central Research.

Harrison, S., Winterbottom, S. & Johnson, R. (2005). *Changing Snow Cover and Winter Tourism and Recreation in the Scottish Highlands*. Hall, C.M. and Higham, J. (eds). Channelview Press, London.

Harvey, S. (2008). *Sicherheit im Bergland - Mustererkennung in der Lawinenkunde*. Österr. Kuratorium für Alpine Sicherheit.

Harvey, S. (2013). Vier Muster für Lawinengefahr. *DAV Panorama*, 1, 56–59.

Harvey, S., Rhyner, H. & Schweizer, J. (2012). *Lawinenkunde - Praxiswissen für Einsteiger und Profis zu Gefahren, Risiken und Strategien*. Bruckmann Verlag GmbH, München.

Harvey, S., Signorell, C. & Genswein, M. (2002). Avalanche Accidents in Back Country Terrain of the Swiss Alps: New Investigations of a 30 year Database. *In: International Snow and Science Workshop, Penticton, British Columbia, Canada*, 449–455.

HIE (1996). *The Economic Impacts of Hillwalking, Mountaineering and Associated Activities in the Highlands and Islands of Scotland*. Highlands and Islands Enterprise.

highland guides (2013). A Historical Comparison in the Development of Scottish Winter Climbing between Ben Nevis and the Northern Cairngorms. *http://www.highlandguides.com/winterhistory.htm [Accessed: 24.04.2013].*

Holden, A. (1998). The Use of Visitor Understanding in Skiing Management and Development Decisions at the Cairngorm Mountains, Scotland. *Tourism Management*, 19 (2), 145–152.

Huberty, C. & Morris, J. (1989). Multivariate Analysis Versus Multiple Univariate Analyses. *Psychological Bulletin*, 105 (2), 302–308.

Jain, A., Duin, R. & Jianchang, M. (2000). Statistical Pattern Recognition: A Review. *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 22 (1), 4–37.

Johnson, R. & Birkeland, K. (2002). Integrating Shear Quality into Stability Test Results. *In: International Snow Science Workshop, Penticton, British Columbia, Canada*, 1–6.

Jolliffe, I. (2002). *Principal Component Analysis*. Springer.

Jomelli, V., Delval, C., Grancher, D., Escande, S., Brunstein, D., Hetu, B., Filion, L. & Pech, P. (2007). Probabilistic Analysis of Recent Snow Avalanche Activity and Weather in the French Alps. *Cold Regions Science and Technology*, 47, 180–192.

Joyce, A. (2001). Weather Variability and Scottish Ice Climbing. *Geographical Journal*, 117 (1), 17–30.

Kienzle, S. (2004). The Effect of DEM Raster Resolution on First Order, Second Order and Compound Terrain Derivatives. *Transactions in GIS*, 8 (1), 83–111.

Kleemayr, K., Moser, A., Schöll, J. & Vollsinger, S. (2000). Lawinenprognose mit statistischen und selbstlernenden Verfahren im Projekt NAFT. *In: Internationales Symposium INTERPRAEVENT 2000 - Villach, Österreich*, 411–421.

Langmuir, E. (1970). Snow profiles in Scotland. *Weather*, 25 (5), 205–209.

Laternser, M. & Schneebeli, M. (2002). Temporal Trend and Spatial Distribution of Avalanche Activity During the Last 50 Years in Switzerland. *Natural Hazards*, 27 (3), 201–230.

lawinenwarndienst tirol (2013). Lawinengefahrenmuster. *http://lawine.tirol.gv.at/basics/lawinen gefahrenmuster/ [Accessed: 19.03.2013].*

Lloyd, C.D. (2010). *Spatial Data Analysis: An Introduction for GIS Users*. Oxford University Press Inc., New York.

Luckmann, B. (1992). Debris Flows and Snow Avalanche Landforms in the Lairig Ghru, Cairngorm Mountains, Scotland. *Geografiska Annaler*, Series A. Physical Geography, 109–121.

MacEachren, A.M., Wachowicz, M., Edsall, R., Haug, D. & Masters, R. (1999). Costructing Knowledge from Multivariate Spatiotemporal Data: Integrating Geographical Visualisation with Knowledge Discovery in Database Methods. *International Journal of Geographical Information Systems*, 13 (4), 311–334.

Maggioni, M. & Gruber, U. (2002). The Influence of Topographic Parameters on Avalanche Release Dimension and Frequency. *In: International Snow Science Workshop, Penticton, British Columbia, Canada*, 97–104.

Mair, R. & Nairz, P. (2010). *Lawine - Die 10 entscheidenden Gefahrenmuster erkennen*. Tyrolia.

Mather, A. (1997). Mountain Recreation in the East Grampians. *Scottish Geographical Magazine*, 113, 195–198.

McCammon, I. (2002). Evidence of Heuristic Traps in Recreational Avalanche Accidents. *In: International Snow Science Workshop, Penticton, British Columbia, Canada*, Citeseer, 244–251.

McCammon, I. (2004). Heuristic Traps in Recreational Avalanche Accidents: Evidence and Implications. *Avalanche News*, 68 (1), 1–10.

McCammon, I. & Hägeli, P. (2007). An Evaluation of Rule-Based Decision Tools for Travel in Avalanche Terrain. *Cold Regions Science and Technology*, 47 (1), 193–206.

McCammon, I. & Schweizer, J. (2002). A Field Method for Identifying Structural Weaknesses in the Snowpack. *In: International Snow Science Workshop, Penticton, British Columbia, Canada*, 477–481.

McClatchey, J. (1996). Spatial and Altitudinal Gradients of Climate in the Cairngorms - Observations from Climatological and Automatic Weather Stations. *Transactions of the Botanical Society*, 48 (1), 31–49.

McClung, D. (2002a). The Elements of Applied Avalanche Forecasting, Part I: The Human Issues. *Natural Hazards*, 26 (2), 111–129.

McClung, D. (2002b). The Elements of Applied Avalanche Forecasting, Part II: The Physical Issues and the Rules of Applied Avalanche Forecasting. *Natural Hazards*, 26 (2), 131–146.

McClung, D. (2003). Magnitude and Frequency of Avalanches in Relation to Terrain and Forest Cover. *Arctic, Antarctic, and Alpine Research*, 35 (1), 82–90.

McClung, D. & Schaerer, P. (1993). *The Avalanche Handbook*. The Mountaineers.

McCollister, C. (2006). Using Geographic Information Systems for Avalanche Work. *The Avalanche Review*, 24 (4).

McCollister, C., Birkeland, K., Hansen, K., Aspinall, R. & Comey, R. (2002). A Probabilistic Technique for Exploring Multi-Scale Spatial Patterns in Historical Avalanche Data by Combining GIS and Meteorological Nearest Neighbors with an Example from the Jackson Hole Ski Area, Wyoming. *In: International Snow Science Workshop, Penticton, British Columbia, Canada*, 109–116.

McCollister, C., Birkeland, K., Hansen, K., Aspinall, R. & Comey, R. (2003). Exploring Multi-Scale Spatial Patterns in Historical Avalanche Data, Jackson Hole Mountain Resort, Wyoming. *Cold Regions Science and Technology*, 37 (3), 299–313.

McCollister, C. (2004). *Geographic Knowledge Discovery Techniques for Exploring Historical Weather and Avalanche Data*. Master's Thesis, Department of Earth Sciences, Montana State University.

McNeally, P. (2008). *Holistic Geographic Visualisation of Spatial Data with Applications in Avalanche Forecasting*. ProQuest.

MetOffice (2013a). Centre for Environmental Data Archival: Met Office SIESAWS. *http://badc.nerc.ac.uk/view/badc.nerc.ac.uk_ATOM_dpt_11694836108612663 [Accessed: 12.06.2013].*

MetOffice (2013b). Mountain Weather Observation. *ww.metoffice.gov.uk/loutdoor/mountainobs/ [Accessed: 17.09.13].*

MetOffice (2013c). Regional Climates. *http://www.metoffice.gov.uk/climate/uk/regional [Accessed: 13.03.2013].*

Miller, H. (2010). The Avalanche Data is Here. *Journal of Regional Science*, 50 (1), 181–201.

Mock, C. & Birkeland, K. (2000). Snow Avalanche Climatology of the Western United States Mountain Ranges. *Bulletin of the American Meteorological Society*, 81 (10), 2367–2392.

Mock, C. & Kay, P. (1992). Avalanche Climatology of the Western United States, with an Emphasis on Alta, Utah. *The Professional Geographer*, 44 (3), 307–318.

Mooi, E. & Sarstedt, M. (2011). Cluster Analysis. *In: A Concise Guide to Market Research*, Springer, Chapter 9, 237–284.

Moss, G. (2009). Avalanche Hazard and Visitor Numbers - A Study in Lochaber, Scotland. *In: International Snow Science Workshop, Davos, Switzerland*, 628–632.

Munros, C.T. (2013). Creag Meagaidh. *http://www.climbthemunros.co.uk/creagmeagaidh.php [Accessed: 20.03.2013].*

Nairz, P. (2008). Lawinen-Muster. *In: Schnee und Lawinen 2007-2008*, Amt der Tiroler Landesregierung, Lawinenwarndienst, Chapter 15, 138–156.

Nairz, P. (2009). Avalanche Patterns as Aid in Avalanche Forecasting. *In: International Snow Science Workshop, Davos, Switzerland*, 389–390.

Nairz, P. & Mair, R. (2010). Was sind Gefahrenmuster? *bergundsteigen*, 4, 62–69.

nevis range (2013). Climbers Information. *http://snowsports.nevisrange.co.uk [Accessed: 13.03.2013].*

Oller, P., Muntán, E., Marturià, J., García, C., García, A. & Martínez, P. (2006). The Avalanche Data in the Catalan Pyrenees. 20 years of Avalanche Mapping. *In: International Snow Science Workshop, Telluride, Colorado, USA*, 305–313.

Pang-Ning, T., Steinbach, M. & Kumar, V. (2006). Cluster Analysis: Basic Concepts and Algorihms. *In: Introduction to Data Mining*, Chapter 8.

Patey, T. (1960). Post-War Winter Mountaineering in Scotland. *The Alpine Journal*, 65 (300/301), 186–194.

Posdnoukhov, A., Matasci, G., Kanevski, M. & Purves, R. (2011). Spatio-Temporal Avalanche Forecasting with Support Vector Machines. *Natural Hazards and Earth System Science*, 11 (2), 367–382.

Posdnoukhov, A., Purves, R. & Kanevski, M. (2008). Applying Machine Learning Methods to Avalanche Forecasting. *Annals of Glaciology*, 49 (1), 107–113.

Price, D. (2000). Scottish Mountain Wind Speeds: A Case-Study at Balquhidder in the Southern Central Highlands. *Weather*, 55, 157–164.

Purves, R., Barton, B., Mackaness, W. & Sugden, D. (1998a). The Development of a Rule-Based Spatial Model of Wind Transport and Deposition of Snow. *Annals of Glaciology*, 26, 197–202.

Purves, R., Morrison, K., Moss, G. & Wright, D. (2003). Nearest Neighbours for Avalanche Forecasting in Scotland - Development, Verification and Optimisation of a Model. *Cold Regions Science and Technology*, 37 (3), 343–355.

Purves, R., Morrison, K. & Wright, B. (1998b). Avalanche Forecasting Experiments for Torridon: A Feasibility Study for sportscotland. *Unpublished Report*.

Purves, R. & Sanderson, M. (1998). A Methodology to Allow Avalanche Forecasting on an Information Retrieval System. *Journal of Documentation*, 54 (2), 198–209.

R (2013). The R Project for Statistical Computing. *http://www.r-project.org/ [Accessed: 20.05.2013].*

Rangachary, N. & Bandyopadhyay, B. (1987). An Analysis of the Synoptic Weather Pattern Associated with Extensive Avalanching in Western Himalaya. *In:* I.P. no 162, ed., *Avalanche Formation, Movement and Effects (Proceedings of the Davos Symposium)*, 311–316.

Roy, M. (1983). The Ben Nevis Meteorological Observatory 1883-1904. *Meteorological Magazine*, 112, 318–329.

SAIS (2013a). Avalanche Map. *http://www.sais.gov.uk/avalanche_map.asp [Accessed: 14.06.2013].*

SAIS (2013b). sportscotland Avalanche Information Service. *http://www.sais.gov.uk/ [Accessed: 11.03.2013].*

Schweizer, J. (2007). Neuer Schneedeckentest - "Nieten" in der Schneedecke suchen. *Die Alpen, Zeitschrift des Schweizer Alpen-Clubs, Bern, Switzerland*, 83 (1), 25–28.

Schweizer, J. (2008). Profilinterpretationen. *In:* B. Zweifel, ed., *Lawinenunfälle in den Schweizer Alpen 2005/2006 - Personen und Sachschäden*, Eidg. Institut für Schnee- und Lawinenforschung SLF, Davos, Switzerland, 129–132.

Schweizer, J. & Jamieson, J. (2003). Snowpack Properties for Snow Profile Interpretation. *Cold Regions Science and Technology*, 37 (7), 233–241.

Schweizer, J. & Jamieson, J. (2010). Snowpack Tests for Assessing Snow-Slope Instability. *Annals of Glaciology*, 51 (54), 187–194.

Schweizer, J., Kronholm, K., Jamieson, J. & Birkeland, K. (2006). Spatial Variability - So What. *In: International Snow and Science Workshop, Telluride, Colorado, USA*, 365–376.

Schweizer, J., Kronholm, K., Jamieson, J. & Birkeland, K. (2008). Review of Spatial Variability of Snowpack Properties and its Importance for Avalanche Formation. *Cold Regions Science and Technology*, 51 (2-3), 253–272.

Schweizer, J. & Lütschg, M. (2000). Measurements of Human-Triggered Avalanches from the Swiss Alps. *In: International Snow Science Workshop, Big Sky, Montana, USA*, 2–6.

Schweizer, J. & Lütschg, M. (2001). Characteristics of Human-Triggered Avalanches. *Cold Regions Science and Technology*, 33 (2), 147–162.

ScottishGovernment (2011). Snow Brings Bumper Winter Sport Season. *http://www.scotland.gov.uk/News/Releases/2011/01/04104338 [Accessed: 19.03.2013].*

Sharp, B. & Whalley, D. (2010). Survey of Scottish Avalanche Incidents (1980-2009). *Report*, The Mountain Rescue Committee of Scotland.

Signorell, C. (2001). *Skifahrerlawinenunfälle in den Schweizer Alpen - Eine Auswertung der letzten 30 Jahre*. Ph.D. Thesis, Universität Basel, Switzerland.

ski mountain (2013). Scottish Ski Conditions and Snow Reports. *http://www.skimountain.co.uk/scottish-ski-conditions [Accessed: 20.04.2013].*

ski scotland (2013). Glencoe. *http://ski.visitscotland.com/centres/glencoe/ [Accessed: 21.03.2013].*

SLF (2013). WSL - Institut für Schnee- und Lawinenforschung SLF. *http://www.slf.ch/ [Accessed: 19.03.2013].*

Steinley, D. (2006). K-means Clustering: A Half-Century Synthesis. *British Journal of Mathematical and Statistical Psychology*, 59 (1-34).

Stoffel, A., Meister, R. & Schweizer, J. (1998). Spatial Characteristics of Avalanche Activity in an Alpine Valley - A GIS Approach. *Annals of Glaciology*, 26, 329–336.

Tase, J. (2004). *Influences on Backcountry Recreationists' Risk of Exposure to Snow Avalanche Hazards*. Master's Thesis, University of Montana.

Thompson, J., Bell, J. & Butler, C. (2001). Digital Elevation Model Resolution: Effects on Terrain Attribute Calculation and Quantitative Soil-Landscape Modeling. *Geoderma*, 100, 67–89.

UKClimbing (2013). Creag Meagaidh. *http://www.ukclimbing.com/logbook/crag.php?id=178 [Accessed: 20.03.2013].*

Vontobel, I. (2011). *Geländeanalysen von Unfalllawinen*. Master's Thesis, University of Zurich.

walkhighlands (2013a). Coire an Lochain. *http://www.walkhighlands.co.uk/cairngorms/coire-an-lochain.shtml [Accessed: 24.03.2013].*

walkhighlands (2013b). Lochnagar. *http://www.walkhighlands.co.uk/munros/lochnagar [Accessed: 13.03.2013].*

Ward, R. (1980). Avalanche Hazard in the Cairngorm Mountains, Scotland. *Journal of Glaciology*, 26 (94), 31–14.

Ward, R. (1981). *Snow Avalanches in Scotland with Particular Reference to the Cairngorm Mountains*. Ph.D. Thesis, University of Aberdeen, Scotland.

Ward, R. (1984a). Avalanche Prediction in Scotland: I. A Survey of Avalanche Activity. *Applied Geography*, 4 (2), 91–108.

Ward, R. (1984b). Avalanche Prediction in Scotland: II. Development of a Predictive Model. *Applied Geography*, 4, 109–133.

Ward, R. (1985). Geomorpholocial Evidence of Avalanche Activity in Scotland. *Geografiska Annaler*, Series A. Physical Geography, 247–256.

Ward, R., Langmuir, E. & Beattie, B. (1985). Snow Profiles and Avalanche Activity in the Cairngorm Mountains, Scotland. *Journal of Glaciology*, 31 (1-7).

Watanabe, S. (1985). *Pattern Recognition: Human and Mechanical*. Wiley, New York.

Watson, A. (1984). Paths and People in the Cairngorms. *Scottish Geographical Magazine*, 100 (3), 151–160.

Wiesinger, T. & Schweizer, J. (2001). Snow Profile Interpretation. *In: International Snow Science Workshop, Big Sky, Montana, USA*, 223–239.

Yuan, M., Buttenfield, B., Gahegan, M. & Miller, H. (2004). *Geospatial Data Mining and Knowledge Discovery*, McMaster, R. and Usery, L.

Zhang, X., Drake, N., Wainwright, J. & Mulligan, M. (1999). Comparison of Slope Estimates from Low Resolution DEMs: Scaling Issues and a Fractal Method for their Solution. *Earch Surface Processes and Landforms*, 24, 763–779.

Appendices

A

APPENDIX

A.1 Data Sources

Table A.1: Overview of the exact data sources of the used data sets.

Data Set		File Format	Organisation	Source	Used Program
Reported Avalanches		.xls	SAIS	Email (Mark Diggins)	Excel / R / ArcGIS
Snow Profiles	**Internet Files**	.csv	SAIS	Webpage SAIS	Excel / R
	CorniceWx	.xls	SAIS	Post (Floyer DVD)	Excel / R
	ProfileObs	.csv	SAIS	Post (Floyer DVD)	Excel / R
	Book1	.xls	SAIS	Email (Dagan Lev)	Excel / R
	Binary Files	.dat	SAIS	Post (Floyer DVD)	Java / Excel / R
Blog		.html	SAIS	Webpage SAIS	Firefox
Annual Reports		.pdf	SAIS	Webpage SAIS	Acrobat Reader
DEM		.asc	University of Zurich	Harddisk	ArcCatalog / ArcGIS
Weather Data	**SIESAWS**	.xls .csv .txt	Met Office	Post (Floyer DVD)	Excel / R
	Cairn Gorm AWS	.csv	Met Office	Webpage	Excel / R

Some additional data sets are stored on the FloyerDVD, but they are not used in this thesis as content and form are not directly relevant to the research questions.

A.2 Detailed Temporal Overview of the Reported Avalanches

Table A.2 provides a detailed overview of the total amount of avalanches that have been reported between 1990 and the end of the season 2012/2013 on a daily resolution.

A.3 Spatial Overview of the Reported Avalanches

Figure A.1 shows an overview of the reported avalanches lying outside of Scotland. Figure A.2 provides an overview of the avalanches per area with different background information including gradient and aspect.

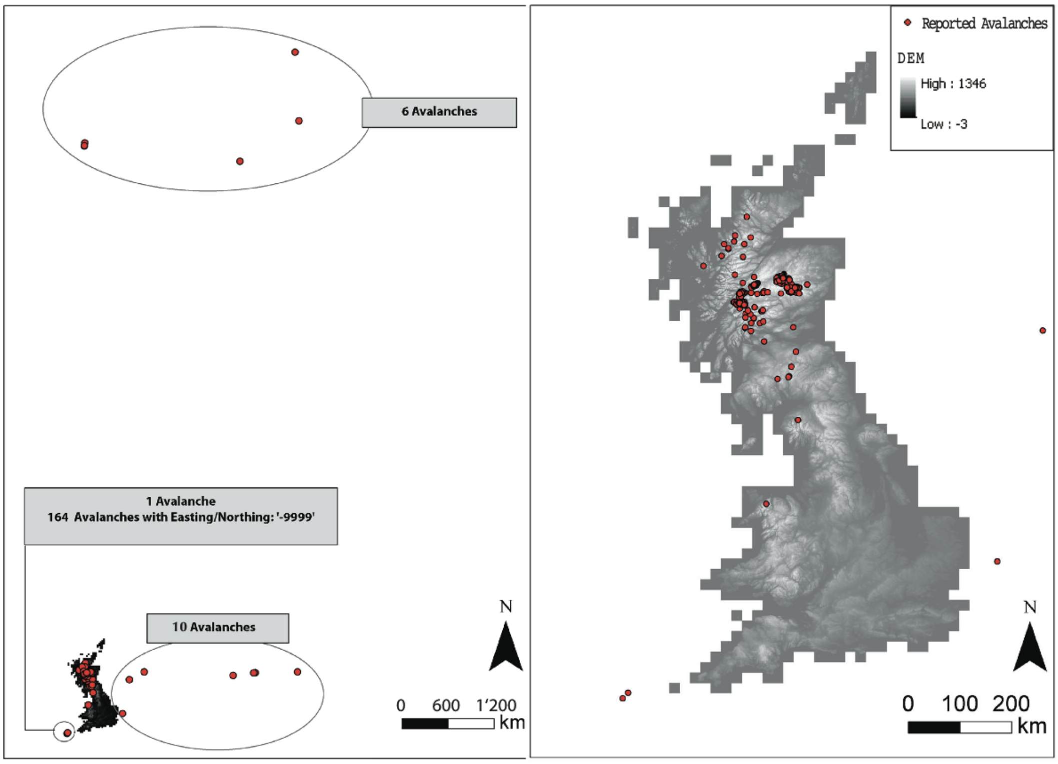


Figure A.1: Left: Locational outliers: avalanches lying outside of Scotland | Right: Total coverage of the DEM for Scotland.

Table A.2: Overview of each avalanche in the reported avalanche data set of the SAIS since 1990 until the end of the season 2012/13.

[illegible]

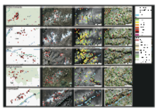


Figure A.2: Overview of the distribution of the reported avalanches (red dots) in the five SAIS areas (background from left to right: Map | Satellite photo | Slope | Aspect).

A.4 Most Important Summits in the SAIS Areas

Figure A.3 shows most important and popular summits in each of the five SAIS areas.

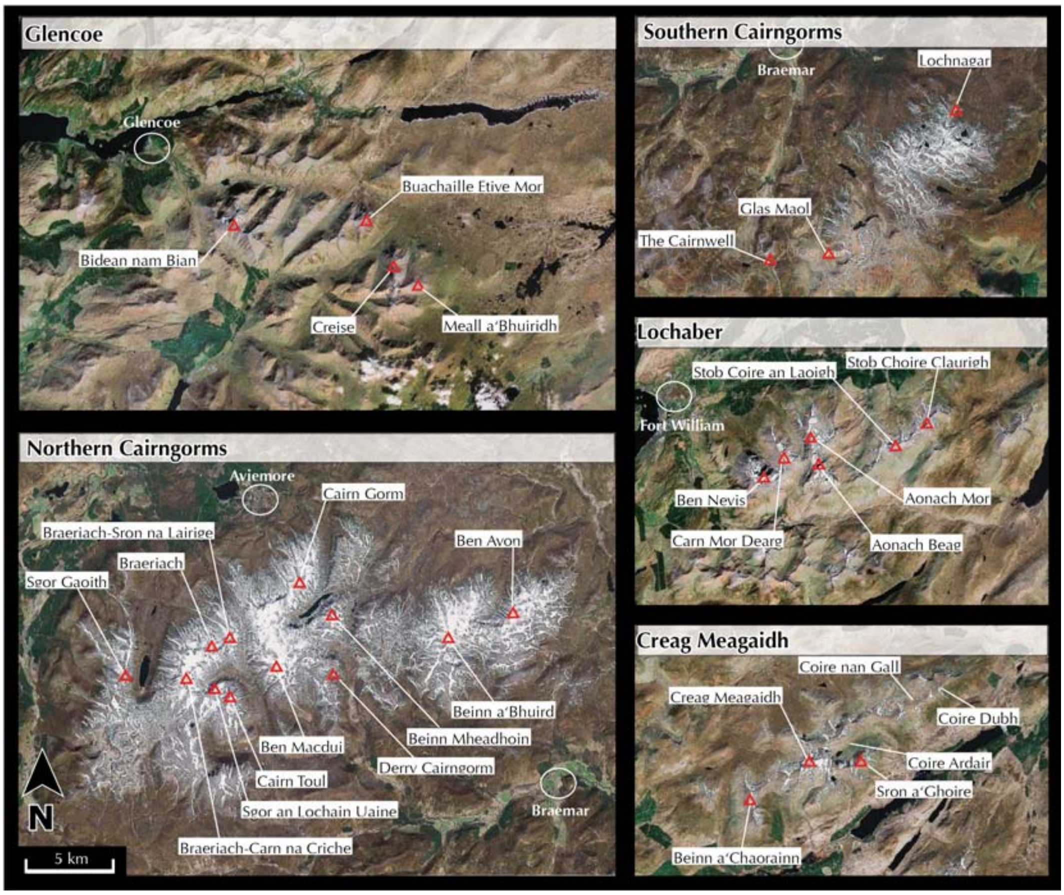


Figure A.3: Overview of each of the five SAIS areas with their most important summits.

A.5 Detailed Methodical Processing Steps

A.5.1 Part A: General Data Analysis

The listing can generally be looked at chronologically, but not without exceptions as the data analysis is an iterative approach.

1) Preparing Data Files To import the data into R or ArcGIS and computing plots, the original data files are pre-processed.

- 'LonLat': divide in two columns 'Longitude' and 'Latitude'.
- 'Date': divide in four columns 'Day', 'Month', 'Year', 'Time'.
- Compute new attributes: grouped 'Daytime', 'Weekday', 'Month' (as text), 'Season', 'Aspect' (as text).
- Search information in 'Comments': search keywords and add new columns: 'Wetness', 'Fracture Type', 'Trigger', 'Number of People', 'Activity', 'Outcome'.
- Identify the meaning of attributes and their values when not clear: 'Type', 'Release', 'Release Attempts', 'Region'.
- Copy of data file and replace all textual NULL values by numerical '-8888' to enable grouped plots.
- Save file as .csv.

2) Computing Frequency Tables and Barplots It is looked at each attribute to identify missing data, typing errors and having a first idea of their distributions.

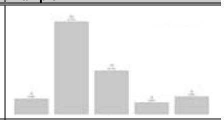
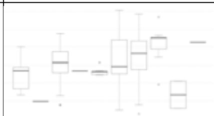
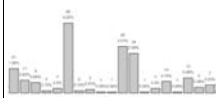
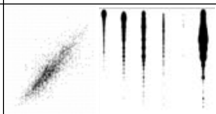
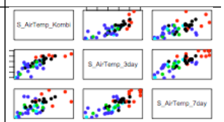
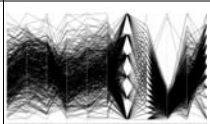
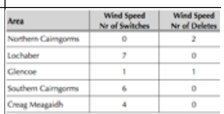
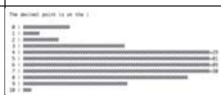
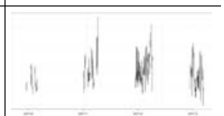
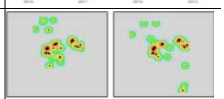
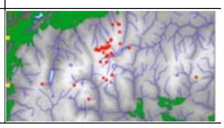
- Read data table into R.
- Compute frequency table for every attribute.
- Look at frequency table and decide which values are NULL values (see table 7.7).
- Look especially at zeros and decide if they are NULL values or have to be used as zero or if some are NULL values and some have to be used as zero (see table 7.7).
- Show data in barplot with different ranges and intervals.
- Look for typing errors (see table 7.7).
- Look at distributions.

3) Computing further Visualisations First impressions of the data or background knowledge give inputs to the assessment of additional aspects of the data using further, often more complex visualisations.

- Plot Weekdays, Months, Years, Seasons per area.
- Plot frequencies: Number of days versus number of avalanches per day (for all areas and in total).
- Comparison of minimum and maximum fracture depth and minimum and maximum width.
- Mapping the points for detection of spatial outliers in ArcGIS.
- Calculate viewshed in ArcGIS.
- Calculate for avalanches lying outside the SAIS areas distances to roads (*'Proximity'* in ArcGIS).
- Calculate and plot point densities for each season (in R).

Possible visualisation techniques mentioned by various authors are summarised in table A.3.

Table A.3: Different visualisation techniques.

Plot	Authors	Classification	Description	Command in R	Example
Barplot	Haining (2003)	univariate	- Shows how many times a value of a certain variable occurrs in a data set.	barplot()	
Boxplot	Fotheringham (2000)	univariate multivariate	- Shows the five most important statistical measures including median, minimum and maximum value and the first and third quartile. - Outliers and extreme values are shown.	boxplot()	
Histogram	Fotheringham (2000) Haining (2003)	univariate	- Shows the distribution of a variable throughout several intervals.	hist()	
Scatterplot	Fotheringham (2000) Haining (2003)	multivariate	- Shows relations between two variables (first variable on x-axis, second on y-axis and each data entry is represented as a point). Can be used as basis for regression analysis and provides a first view of clusters, outliers or trends. - To account for multiple values at coordinate, the point size can be adjusted (Bubbleplot) or different symbols can be used. - Also 3D plots including three variables are possible.	plot() sizeplot()	
Scatterplot Matrix	Fotheringham (2000) Haining (2003)	multivarirate	- Scatterplots of all pairs of variables in a data set. - Grouped objects can be plotted with different colors to get an impression of clusters.	pairs() plot()	
Parallel Coordinate Plots	McEachren (1999) Fotheringham (2000) Haining (2003)	multivariate	- Each entry of the data set represents a horizontal line above the x-axis which consists of all attributes. - Gives a first impression of relationships between variables.	parcoord()	
Statistical Measures	Haining (2003)	univariate	- Different measures are possible: for example Mean, Median, Standard Deviation, Variance, Minimum and Maximum Values	summary() describe()	Min. 1st Qu. Median Mean 3rd Qu. Max. 1.000 3.000 4.000 4.062 5.000 9.000
Tables		univariate multivariate	- Summarising data. - Textual / numerical description of certain properties of variables. - Frequency tables are possible.	table()	 Area / Wind Speed Nr of Switches / Wind Speed Nr of Deletes Northern Cairngorms / 0 / 2 Lochaber / 7 / 0 Glencoe / 1 / 1 Southern Cairngorms / 6 / 0 Creag Meagaidh / 4 / 0
Stem and Leaf Plot	Fotheringham (2000) Haining (2003)	univariate	- Near-graphical view of the distribution of the data. - Information of each data value is still included.	stem()	
Time Series Plots	Haining (2003)	univariate multivariate	- Shows how variables change over time. - The time is shown on the x-axis and one or more variables on the y-axis.	plot()	
Density Estimates	Fotheringham (2000)	univariate	- Density functions are used to calculate densities - Examples include Kernel density - Can be plotted spatially or with the density curve of a variable	smoothScatter()	
Maps	Fotheringham (2000)	univariate	- Spatial data can be plotted - Variable can be viewed according to their spatial distributions. - Examples are choropleth maps or cartograms.	plot() points()	

A.5.2 Part B: Data Pre-processing

The Binary data is stored on the FloyerDVD in about 180 different folders and sub folders without big systematic (see figure A.4 for an example). The filenames consist of a shortening of the area, year, month and day. Some files are stored in several folders in different combinations, sometimes more than once, sometimes with different filenames, sometimes with day and month mixed up or with a wrong area abbreviation. In a first step, these binary files are manually selected and combined in one folder per area, so that a systematic, ordered basic data set is available which is as complete and consistent as possible.

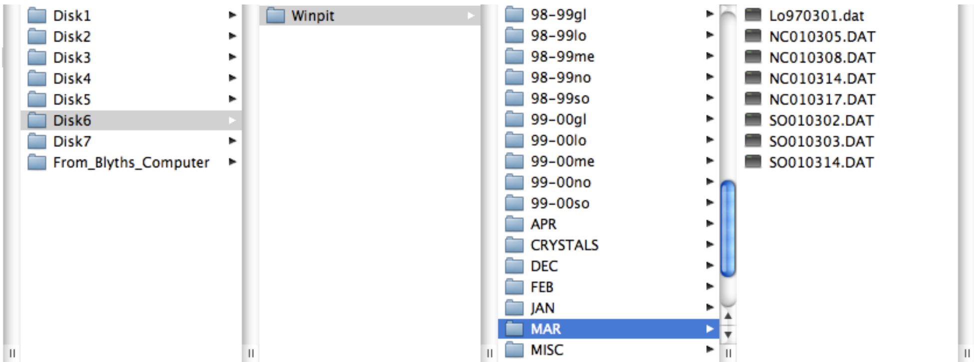


Figure A.4: Example of original binary file data structure.

The files are read using a Java program. The main functionality was already written by James Floyer, but the program is adapted to enable further analysis. Not only one but all files can be read and the header and depth information is treated and stored separately in a comma separated file for each single area (see appendix A.6 for code).

The Internet data is downloaded from the webpage of the SAIS (SAIS 2013b: Snow Profiles) as .csv files which are imported into Excel, split up into the five SAIS areas and an area id[1] is inserted.

The ProfileObs data is stored on the FloyerDVD in two different .csv files for each area (one up until 2003 and one from 2007 to 2010). The date is changed into the consistent format *DD.MM.YYYY* from *DD-MonthText-YYYY* (before 2003) and *DD/MM/YYYY* (from 2007 to 2010). The area names are formatted in a consistent way and the area id added. The observed and forecasted avalanche hazard are changed from text format into numbers ('Low' = 1, 'Moderate' = 2, 'Considerable' = 3, 'High' = 4).

The CorniceWx data is stored on the FloyerDVD as one .xlsx file per area. The date is changed into the format *DD.MM.YYYY* from *DD/MM/YYYY*. The area id is inserted. Columns with values from the Northern Cairngorms in the file of the Southern Cairngorms are deleted.

The Book1 data is only used for the depth information (see appendix A.6.2).

[1]The area id is always chosen the following: NC: 2, LO: 3, GL: 4, SC: 5, ME: 6.

A.6 Java Code to Read the Binary Files

```java
import java.io.*;
import java.util.*;

/* James Floyer worked this bit out!
'following attributes are stored in the header of the data files:
'Field name             No. chars       Format                  Data checks
'Air Temp                   6           Float                   numeric; >-50 and <50
'Altitude                   4           Integer                 numeric; >=0 and <1500
'Region                     20          String
'Aspect                     3           Integer                 numeric; 0 to 360; if 360 then =0
'Avalanche Code             5           String                  numeric
'Time                       4           DateTime                numeric
'Cloud                      3           Integer                 numeric; 0 to 100
'Comment                    85          String
'Date                       10          DateTime                DateTime format; 8-digit dates
                                                               (dd/mm/yy) converted to yyyy/mm/dd
'Drift Y/N                  1           String                  y, Y, n, N; output capitilized
'Foot Pen                   3           Integer                 numeric; >=0 and <100
'Forecast Avalanche Hazard  2           String                  1, 2, 3, 3-, 3+, 4, 5
'Grid Ref                   6           String                  numeric
'Incline                    2           Integer                 numeric; 0 to 90
'Location                   85          String
'Observer                   14          String
'Observed Avalanche Hazard  3           String                  1, 2, 3, 3-, 3+, 4, 5
'Precipitation Code         4           Integer                 numeric; 0 to 10; need to check legitimate codes
'Ski Pen                    3           Integer                 numeric; >=0 and <100
'Total Snow Depth           4           Integer                 numeric; >=0 and <1000
'Wind Direction             3           Integer                 numeric; 0 to 360; if 360 then =0
'Wind Speed                 3           Integer                 numeric; >=0

following attributes are stored in the depth information of the data files:
'Field name             Format
'Comments               String
'Size (E)               String
'Form (F)               String
'Form1 (F)              String
'Unknown Code           String
'Comment Depth          Float
'Temperature            Float
'Temperature Depth      Float
'Layer Depth            Float
'Wetness                Float
*/


public class ReadWinPit_total{

        //Define Separator to ensure an import into Excel
        private final String SEP = "$";

        //Define Region for which the files should be read and stored
        //String region = "NC"; //Northern Cairngorms
        //String region = "LO"; //Lochaber
        //String region = "GL"; //Glencoe
        //String region = "SC"; //Southern Cairngorms
        //String region = "ME"; //Creag Meagaidh
        String region = "";   //total, all regions

        //Define Source Folder containing the Files which have to be read
        private final String sourceFolder = "/SP_Files/"+region;
        //Define Folder and Name of the resulting File for the Header Information
        private final String resultFileHeader = "/allHeader_"+region+".txt";
        //Define Folder and Name of the resulting File for the Depth Information
        private final String resultFileDepthInfo = "/allDepthInfo_"+region+".txt";

        //Defines the Number of Chars used in the binary data file.
        private static int [] lengths = {6,4,20,3,5,4,3,85,10,1,3,2,6,2,85,14,3,4,3,4,3,3};

        /**
         * Main Method of the Snow Profile Data Reading Class
         * @param args
         * @throws Exception
         */
        public static void main(String args[]) throws Exception{
                new ReadWinPit_total().go();
        }
```

```java
        /**
         * Chief method.
         * Defining of the data to read
         * Preparing list with all files
         * Run through each file and read the file and make new line
         * Write to new file
         * @throws Exception
         */
        void go() throws Exception {
                //Invokes method which reads the header of the binary files and writes it to a file.
                goHeader();
                //Invokes method which reads the depth information of the binary files and writes it to a file.
                goDepthInfo();
        }

        /**
         * Gets the list of the files
         * Goes through each file and append the information form the header to the text
         * Writes the text to a file
         * @throws Exception
         */
        void goHeader() throws Exception{
                String fileName = sourceFolder;
                        ArrayList<File> aList = walk(fileName);
                        StringBuilder sb = new StringBuilder();
                        for (int i=0; i<aList.size(); i++) {
                                sb.append(readHeader(aList.get(i).getAbsolutePath()));
                                sb.append("\n");
                        }
                        //write data to text file;
                        writeFile(new File(resultFileHeader), sb.toString(), false);
        }

        /**
         * Gets the list of the files
         * Goes through each file and append the depth information to the text
         * Writes the text to a file
         * @throws Exception
         */
        void goDepthInfo() throws Exception{
                String fileName = sourceFolder;
                        ArrayList<File> aList = walk(fileName);
                        StringBuilder sb = new StringBuilder();
                        for (int i=0; i<aList.size(); i++) {
                                sb.append(readDepthInfo(aList.get(i).getAbsolutePath()));
                        //sb.append("\n");
                        }
                        //write data to text file;
                        writeFile(new File(resultFileDepthInfo), sb.toString(), false);
        }

        /**
         * Reads the header information of a snow profile file and stores it in one string
         * @param fileName
         * @return
         * @throws Exception
         */
        private String readHeader(String fileName) throws Exception {
                StringBuilder sb = new StringBuilder();
                BufferedReader br = new BufferedReader(new FileReader(fileName));
                int pos = 0;
                //Read the header and store into StringBuilder
                for (int i=0; i < lengths.length; i++){
                        char input[] = new char[lengths[i]];
                        pos = pos + br.read(input, 0 , lengths[i]);
                        sb.append(String.valueOf(input).trim()+SEP);
                }
                return sb.toString();
        }
```

Figure A.5: Java code to read the binary files, part one.

```java
/**
 * Reads the depth information of a snow profile file and stores it in one string
 * @param fileName
 * @return
 * @throws Exception
 */
private String readDepthInfo(String fileName) throws Exception{

        StringBuilder sb = new StringBuilder();
        BufferedReader br = new BufferedReader(new FileReader(fileName));
        int pos = 0; //Initialize variable which stores the position of information in the file

        //Preparing arrays for each attribute
        String[] Comments = new String[126];   //Comments containing additional information in text form
                                               //mostly contains info about Walking Rutschblock Test (WRB),
                                                    Shovel Shear Test (SST) ect.
        String[] Size = new String[126];       //Describes the size of the snow crystals in [cm]
        String[] Form = new String[126];       //Describes the form of the snow crystals
                                               //1:Precipitation particles, 2:Decomposing/fragmented particles,
                                                  3:Rounded grains, 4:Faceted crystals, 5:Depth hoar, 6:Wet grains,
                                                  7:Feathery crystals/surface hoar, 8:Ice masses, 9:Surface
                                                  deposits and crusts
        String[] Form1 = new String[126];      //Describes the form of the snow crystals, interpretation see Form
        Float[] UnknownCode = new Float[126];
        Float[] CommentDepth = new Float[126]; //Depth of the profile to which the Comments belong to
        Float[] Temp = new Float[126];         //Measured temperature in [°C]
        Float[] TempDepth = new Float[126];    //Depth of the profile to which the temperatures belong to
        Float[] LayerDepth = new Float[126];   //Depth of the layers which occur in the profiles
        Float[] Wetness = new Float[136];      //Wetness of the layer
                                               //1:No snowball, 2:Snowball, 3:Wet glove, 4:Water drops, 5:Slush
        String region = new String();
        String date = new String();
        String comment = new String();

        //Reading the header information
        for (int i=0; i < lengths.length; i++){
                char input[] = new char[lengths[i]];
                pos = pos + br.read(input, 0 , lengths[i]);
                //Storing only the Region (at pos30), the Data (at pos140) and the first Comment (at pos313)
                to the StringBuilder
                //if(pos==30 || pos==140 || pos==313){
                //append Region, Date and first Comment to the StringBuilder
                //sb.append(String.valueOf(input).trim()+SEP);
                //}
                if(pos==30) region=String.valueOf(input).trim();
                else if(pos==140) date = String.valueOf(input).trim();
                else if(pos==313) comment = String.valueOf(input).trim();
        }

        //Reading Comments (up to nine are possible) and storing into array
        for (int i=0; i < 9; i++){
                char input[] = new char[40];
                pos = pos + br.read(input, 0 , 40);
                //store additional comments into array
                Comments[i] = String.valueOf(input).trim();
        }

        //Jump to location of the size information in the file
        int skip = 1260 - pos;
        br.skip(skip);

        //Reading Size and storing into array
        for (int i=0; i < 30; i++){
                char input2[] = new char[7];
                pos = pos + br.read(input2, 0 , 7);
                //store sizes into array
                Size[i] = String.valueOf(input2).trim();
        }

        //Reading Form and storing into array
        for (int i=0; i < 30; i++){
                char input3[] = new char[2];
                pos = pos + br.read(input3, 0 , 2);
                //store forms into array
                Form[i] = String.valueOf(input3).trim();
        }

        //Reading Form1 and storing into array
        for (int i=0; i < 30; i++){
                char input4[] = new char[2];
                pos = pos + br.read(input4, 0 , 2);
                //store form1s into array
                Form1[i] = String.valueOf(input4).trim(); }

        //Store content of file as DataInputStream
        DataInputStream dis = new DataInputStream(new FileInputStream(new File(fileName)));

        //Jump to location of Depth information in the file
        dis.skip(677);

        //Initialize help variables
        int i=0;
        int count =0;

        //Reading Unknown Code and storing into array
        for (i=i; i < 10; i++){
                UnknownCode[i]= ReadWinPit_total.readFloat(dis);
                count++; }

        count =0;
        //Reading Depths of the Comments and storing into array
        for (i=i; i <= 20; i++){
                CommentDepth[i]= ReadWinPit_total.readFloat(dis);
                count++; }

        count =0;
        //Reading Temperatures and storing into array
        for (i=i; i <= 51; i++){
                Temp[i]= ReadWinPit_total.readFloat(dis);
                count++; }

        count =0;
        //Reading Depths of Temperatures and storing into array
        for (i=i; i <= 81; i++){
                TempDepth[i]= ReadWinPit_total.readFloat(dis);
                count++; }

        count =0;
        //Reading Depths of Layers and storing into array
        for (i=i; i <= 114; i++){
                LayerDepth[i]= ReadWinPit_total.readFloat(dis);
                count++; }

        count =0;
        //Reading Wetnesses and storing into array
        for (i=i; i <= 125; i++){
                Wetness[i]= ReadWinPit_total.readFloat(dis);
                count++; }

        //append the first 20 values of each attribute to the Stringbuilder
        //separately for each layer depth
        //order: Comment, CommentDepth, Size, Form, Form1, Wetness, Layer Depth, Temperature, Temperature Depth
        //add a separator between each value to ensure a error free reading into excel
        int id = 1;
                for(int j=0; j<20; j++){
                        sb.append(region+SEP+date+SEP);
                        sb.append(id);
                        sb.append(comment+SEP);
                        sb.append(Comments[j]+SEP);
                        sb.append(CommentDepth[10+j]+SEP);
                        sb.append(Size[j]+SEP);
                        sb.append(Form[j]+SEP);
                        sb.append(Form1[j]+SEP);
                        sb.append(Wetness[115+j]+SEP);
                        sb.append(LayerDepth[82+j]+SEP);
                        sb.append(Temp[21+j]+SEP);
                        sb.append(TempDepth[52+j]+SEP);
                        sb.append("\n"); //new line
                        id++;
                }

                //change the Stringbuilder to String and return the String containing the depth information
                for one snow profile
                return sb.toString();
}
```

Figure A.6: Java code to read the binary files, part two.

```java
// From http://www.peterfranza.com/2008/09/26/little-endian-input-stream/
public static int readInt(DataInputStream d) throws Exception{
        byte[] w = new byte[8];
        ReadWinPit_total.readFully(w, 0, 4, d);
        return
        (w[3])      << 24 |
        (w[2]&0xff) << 16 |
        (w[1]&0xff) <<  8 |
        (w[0]&0xff);
}

public static float readFloat(DataInputStream d) throws Exception{
        return Float.intBitsToFloat(readInt(d));
}

public static void readFully(byte b[], int off, int len, DataInputStream d) throws Exception{
        d.readFully(b, off, len);
}

/**
 * Goes through a directory and puts all files into an array list
 * @param path
 * @return ArrayList
 */
public ArrayList<File> walk(String path) {

       ArrayList<File> aList = new ArrayList<File>();
       File root = new File(path);
       File[] list = root.listFiles();

       for (File f : list) {
           if (f.isDirectory()) {
               aList.addAll(walk( f.getAbsolutePath() ));
           }
           else {
               if (f.getName().endsWith(".DAT")){
                   aList.add(f.getAbsoluteFile());
               }
           }
       }
       return aList;
   }

/**
 * Writes a text to a file
 * @param file
 * @param text
 * @param append
 */
public void writeFile(File file, String text, boolean append) {
         try {
                FileWriter writer = new FileWriter(file, append);
            writer.append(text);
            writer.close();
         }
         catch (IOException e) {
            e.printStackTrace();
         }
}
}
```

Figure A.7: Java code to read the binary files, part three.

A.6.1 Part B: Combination of Header Information

Textual values for the 'Observed Avalanche Hazard' and 'Forecasted Avalanche Hazard' ('Low', 'Moderate', 'Considerable', 'High', 'Very high') as well as for 'Drift' ('Y', 'N') are replaced in the *internet files* by the corresponding numbers 1, 2, 3, 4, 5 for Avalanche Hazards and 1 or 0 for Drift.

3- and 3+ values in the attributes 'Observed Avalanche Hazard' and 'Forecasted Avalanche Hazard' in the *ProfileObs* data are changed to 3.

'Air Temperature' values for the Northern Cairngorms, Glencoe and Creag Meagaidh are processed further to reach consistency. The values in the *internet files* are sometimes given in 1/10 °C and sometimes in °C (see section 7.2.2.3 for details).

'Wind Speed' and 'Summit Wind Speed' values seemed to be exchanged with 'Wind Direction' and 'Summit Wind Direction' values in some cases (see section 7.2.2.3 for details).

'Precipitation Code' and 'Snow Index' values are combined, as they have the same content but sometimes different temporal coverages. Snow Index values are used if they are present, otherwise the Precipitation Code values are taken.

A.6.2 Part B: Combination of Depth Information

- New columns for 'Time', 'ID', 'ProfileID' and 'Hardness' are added to the *binary file* data.
- New columns for 'CommentDepth' and 'Form1' are added to the *Book1* data in the sheet *'ProfileLayers'*.
- Empty rows are inserted to every profile in *Book1* data in both sheets using a VBA script in Excel so that every profile consists of 20 rows to enable the combination.
- Layer id from 1 to 20 is added to each profile in both sheets of *Book1*.

- Temperature and temperature depth from *Book1* sheet *’TempLayer’* are copied to *Book1* sheet *’ProfileLayers’*.
- Data from *Book1* are copied to the data of the *binary files*.
- Some steps to ensure consistency are processed: ’:’ in some of the values in the attribute ’Form’ and ’Form1’ are deleted. ’Size’ values such as *00.5-1.*, *50.5-0.* or *0>0.5* are manually changed to most sensible next bigger numbers.
- Empty rows are deleted.

A.6.3 Part B: Temporal Coverage of Snow Profiles and Avalanches

Dates on which avalanches released but no snow profiles are available: 5.12.2009 (3), 6.11.2010 (2), 14.11.2010 (4), 24.11.2010 (1), 25.11.2010 (1), 10.4.2011 (1), 3.5.2011 (1), 2.12.2011 (2), 3.12.2011 (1), 5.12.2011 (1), 7.12.2011 (1), 12.12.2011 (2), 13.12.2011 (1), 29.4.2012 (1), 2.5.2012 (1), 7.5.2012 (1), 20.5.2012 (1), 3.11.2012 (2), 5.11.2012 (1), 6.11.2012 (1), 28.11.2012 (1), 6.12.2012 (1), 9.12.2012 (1).

A.6.4 Part B: Temporal Coverage of the Binary Files

The data of the *binary files* is shown very generalised in table 7.11, as sometimes only few days in a certain area are missing. A more detailed temporal coverage of the *binary files* is provided in table A.4.

Table A.4: Number of binary files available per season and month for each area.

Season	NC	LO	GL	SC	ME
89-90	**37**				
Dez					
Jan	28				
Feb	7				
Mar					
Apr					
90-91			**128**		
Dez			18		
Jan			31		
Feb			28		
Mar			31		
Apr			20		
91-92			**122**		
Dez			12		
Jan			31		
Feb			29		
Mar			31		
Apr			19		
92-93		**47**	**118**		
Dez			14		
Jan			31		
Feb		2	28		
Mar		31	31		
Apr		14	14		
93-94	**67**	**117**	**10**		
Dez		12	10		
Jan		31			
Feb	21	28			
Mar	31	31			
Apr	15	15			
94-95	**2**	**51**			
Dez	1	14			
Jan		31			
Feb		6			
Mar	1				
95-96	**119**	**118**	**118**		
Dez	15	14	14		
Jan	31	30	31		
Feb	29	29	29		
Mar	31	31	31		
Apr	13	13	13		

Season	NC	LO	GL	SC	ME
96-97	**112**	**111**	**112**	**107**	**115**
Dez	16	16	16	15	19
Jan	31	30	31	31	31
Feb	28	28	28	26	28
Mar	31	31	31	2	31
Apr	6	6	6	6	6
97-98	**117**	**118**	**119**	**116**	**119**
Dez	15	16	17	19	17
Jan	31	31	31	31	31
Feb	28	28	28	28	28
Mar	31	31	31	29	31
Apr	12	12	12	9	12
98-99	**117**	**117**	**115**	**117**	**117**
Dez	16	16	16	16	16
Jan	31	31	30	31	31
Feb	28	28	28	28	28
Mar	31	31	31	31	31
Apr	11	11	10	11	11
99-00	**107**	**104**	**103**	**103**	**109**
Dez	7	9	9	9	10
Jan	31	31	31	31	31
Feb	29	29	29	29	29
Mar	31	31	31	31	31
Apr	9		3	3	8
00-01	**116**	**103**	**97**	**106**	**93**
Nov	4				
Dez	12	12	12	12	12
Jan	31	31	31	31	31
Feb	28	28	28	27	26
Mar	25	17	11	21	9
Apr	16	15	15	15	15
01-02	**95**	**101**	**101**	**100**	**101**
Dez	11	11	11	11	11
Jan	31	31	31	30	31
Feb	26	28	28	28	28
Mar	27	31	31	31	31

Season	NC	LO	GL	SC	ME
02-03	**24**	**122**	**91**	**122**	**115**
Dez	16	16	17	17	16
Jan	6	31	31	31	31
Feb	1	28	28	28	28
Mar	1	31		31	31
Apr		15	15	15	9
May		1			
03-04		**122**	**78**		**122**
Dez		18	17		17
Jan		31	31		31
Feb		29	29		29
Mar		30	1		31
Apr		14			14
04-05	**11**	**122**			**109**
Dez	11	17			11
Jan		31			30
Feb		28			28
Mar		31			31
Apr		15			9
05-06		**71**			**109**
Dez		17			10
Jan		31			31
Feb		23			28
Mar					31
Apr					9
06-07					**107**
Dez					10
Jan					31
Feb					28
Mar					7
Apr					

A.6.5 Parts A and B: Data Quality Analysis

Detailed data quality aspects that are analysed for the data set of the reported avalanches (Part A, signed with [AVAL] and of the snow profiles (Part B, signed with [SP]):

Accuracy:

- ***Measurement imprecision / Spatial object***: locational outliers are detected by mapping the avalanche points. [AVAL]
- ***Measurement imprecision / Variable value***: values of attributes for which other data sources are available are compared.
  - [AVAL]: the attribute 'Region' is plotted and compared with a map. The attributes 'Altitude', 'Aspect' and 'Gradient' are compared with the values extracted from the DEM using scatterplots.
  - [SP]: the related attributes 'Air Temperature', 'Wind Direction' and 'Wind Speed' measured at the snow profile locations and at the summit stations are compared using scatterplots.
- ***Gross errors / variable value***: typing errors in each attribute are extracted by looking at the distribution of the attribute values in barplots. [AVAL & SP]

Resolution:

- ***Spatial resolution***: makes no sense for avalanches, but is assessed for the DEM and the snow profiles by mapping and considering the task. [DEM & SP]
- ***Temporal resolution***: is analysed for the snow profile attributes by analysing their units, but makes no sense in the avalanche data as no time series are included. [SP]
- ***Variable resolution***: assessed for each attribute using barplots. [AVAL & SP]

Consistency:

- ***Internal***: assessed for each attribute using barplots. [AVAL & SP]
- ***Internal***: different labels for missing data are analysed using barplots. [AVAL & SP]
- ***Internal***: related attributes are compared and the relations checked on sensibility.
  - [AVAL]: values of the attributes 'Minimum Fracture Depth' and 'Maximum Fracture Depth' as well as 'Minimum Width' and 'Maximum Width' are compared by analysing the data matrix.
  - [SP]: the attributes 'Wind Direction' and 'Wind Speed' as well as 'Summit Wind Direction' and 'Summit Wind Speed' are compared by looking at maximum values.
- ***Linking data sets***: identical attributes in different data sources are compared in the snow profile data set, especially the summit measures 'Summit Air Temperature', 'Summit Wind Speed' and 'Summit Wind Direction' using scatterplots. For the avalanche data set, it is not necessary as only one data set is available. [SP]

Completeness:

- ***Spatial***: the spatial coverages of the DEM and the snow profiles are analysed by mapping.
- ***Variable values***: the amount of missing data is analysed for each attribute using frequency tables. [AVAL & SP]
- ***Variable values***: the amount of the original data is compared with the amount of the enhanced data set for the attributes 'Region', 'Altitude', 'Aspect' and 'Gradient' by analysing the data matrix and using frequency tables. [AVAL]
- **+ *Temporal coverage***: time spans covered by each attribute of the snow profile data sets are assessed for each area separately by analysing the missing data within the attribute 'Date'. [SP]

Model quality

- ***Spatial objects***: the question 'Are points an accurate representation of avalanches?' is answered applying background knowledge and considering the current tasks. [AVAL]
- ***Spatial relationships***: as only points are analysed spatial relationships and topology rules are not important. [AVAL & SP]
- ***Attributes***: the question 'Are all necessary attributes included in the data set?' is answered by comparing the available attributes with the background knowledge and other studies. [AVAL & SP]

A.6.6 Part C: Steps of PCA

1. Preparation of the data, so that variables and objects are obtained in either rows or columns. It has to be decided on which variables are included in the input and how will be dealt with missing values. The data has to be standardised by scaling and centering to avoid a diverse weighting of variables. This can be done by subtracting the mean of the variable from each value and divide each value through the standard deviation. In this way, the new mean equals zero and the standard deviation equals one (Backhaus *et al.* 2008)
   R: *na.omit(), scale()*
2. A correlation matrix has to be calculated showing the correlation coefficients between every pair of variables. Correlation coefficients range between 0 and 1 and the bigger they are the higher is the correlation. These values give an indication of variables which are related to each other but it says nothing about the direction of determination or the number of factors that have an influence (Backhaus *et al.* 2008)
   R: *cor(Matrix)*
3. Several steps can be done for looking how well the data is suitable for a PCA. It can be looked at the ***correlation coefficients***, ***test of normality*** can be applied and ***inverse correlation matrix***, ***Bartlett-Test***, ***Anti-Image -Kovariance-Matrix*** or ***Kaiser-Meyer-Olkin-Criterium*** can be used (Backhaus *et al.* 2008).
   R: *cor.test(), solve()*
4. The PCA can be run the first time using all variables. As results the following four parameters are calculated:
   A. **Factor values**: describe the amount on which a component is influenced by the variable. Positive values mean over average influence, zero means average influence and negative values stand for less than average influence.
   B. **Factor loadings**: describe the amount on which the variables correlate with the components.
   C. **Eigen vectors**: are the linear combinations of the variables which make up a component.
   D. **Eigen values**: describe the part of the total variability that is explained by the component. If the values are smaller than 1, the component explains less of the total variability as his internal variability is.

   R: *prcomp(), pca(), princomp(), svd(), PCA()*
5. Mostly the next step includes the decision on the number of components to retain for the further steps. The fewer components are used, the more information gets lost, the more components are used, the more information can be maintained but the less data reduction can be provided. This decision can be done using the ***Kaiser-Criteria***, ***Scree-Test*** or ***amount of total variation*** that is explained by the components (Jolliffe 2002, Backhaus *et al.* 2008).
   R: *screeplot() and lines(), barplot()*
6. To obtain a clear assignment of the components a rotation can be applied on the data.
   R: *varimax()*
7. The results can be tested and interpreted using various graphical representations, for example biplots (Backhaus *et al.* 2008)
   R: *plot(), biplot(), dotplot()*

A.6.7 Part C: Steps of CA

1. Preparation of the data, so that the variables and objects are obtained in either rows or columns. It has to be decided on which variables are included in the input and how will be dealt with missing values. The data has to be standardised, so scaled and centred to avoid a different weighting of variables. This can be done by subtracting the mean of the variable from each value and divide each value through the standard deviation (Backhaus *et al.* 2008).
   R: *na.omit(), scale(Data, center=TRUE, scale=TRUE)*
2. A proximity matrix is calculated using either a ***similarity*** or a ***dissimilarity*** measure. Further can be distinguished between distance measures where the absolute values are important or similarity measures in which the relative trend in the data is significant. The choice is done depending on the scale of the data. For metric data can the ***city-block metric***, ***Pearson Correlation coefficient*** or ***Euclidean Distance*** (squared, summed and rooted values) be used (Backhaus *et al.* 2008).
   R: *dist(Data, method=m) with m: 'euclidean', 'maximum', 'manhattan', 'canberra', 'binary' or 'minkowski'*
3. The clustering algorithm has to be chosen, where mostly ***partitioning*** and ***hierarchical*** methods are distinguished. With partitioning methods a predefined number and mostly random partition of cluster is set and the objects are then reallocated and exchanged between the clusters in order to reach an optimized variance (squared sum of errors to the mean). Partitioning methods include for example the popular ***k-means*** method. It does not use a distance measure but assigns the objects in such a way to the clusters that the within-cluster variation (sum of squared distances to the geometric cluster centre) is minimised. An iteration of calculating the Euclidean distances to cluster centres, reallocate objects and calculating new cluster centres is done.
   Hierarchical methods can further be divided into ***agglomerative*** (bottom-up) and ***divisive*** (top-down) ones. The first approach starts with each object being one single cluster which are then merged in several steps until only one cluster including all objects is present. Methods for merging clusters include ***single linkage*** ("Nearest Neighbour"-method, uses always the smallest distance between single objects, tends to build many, small groups, good for detecting outliers), ***complete linkage*** ("Furthest Neighbour"-method, uses always the largest distance between single objects, tends to build smaller groups, affected by outliers), ***average linkage*** (uses the average distance between all pairs of objects in the clusters, tends to build groups with lower variance and similar sizes), ***centroid*** (uses the distance between the geometric centre of the clusters, tends to build groups with lower variance and similar sizes), ***median*** or ***Ward*** (uses the variance as measure of heterogeneity and merges those objects which increase the variance in a cluster the least). The second approach starts with one cluster which includes all objects and is then divided stepwise until each cluster consists of only one object (Backhaus *et al.* 2008).
   R: *hclust() and cutree(Data.hclust, k) with k: nr of clusters, kmeans(Data, k) with k: nr of clusters*
4. The number of clusters can be defined using the ***dendrogram***, ***Scree plots***, the ***criteria of Calinski and Harabasz*** or the ***test of Mojena*** (Backhaus *et al.* 2008)
   R: *plot(Data.hclust) and rect.hclust(Data.hclust)*
5. To interpret the resulting clusters for example ***F-Values***, ***t-values*** or a ***discriminant analysis*** can be used (Backhaus *et al.* 2008). It can be looked at the ***cluster centroids*** and tested if they are distinguishable with independent t-tests or ANOVA (Mooi & Sarstedt 2011)
6. To validate the results the stability can be assessed by changing the ***proximity measure***, the ***algorithm*** or the ***number of clusters*** and comparing different results (Backhaus *et al.* 2008). Otherwise, the effect of ***splitting the original data matrix*** or ***reordering of the objects*** in the original data matrix can be assessed (Mooi & Sarstedt 2011).

A.7 Content of the DVD Including Additional Material

Thesis:

- ***MasterThesis_EvelineBraun_2013.pdf***: Official thesis document.
- ***MasterThesis_EvelineBraun_2013.rtf***: Official thesis document as .rtf
- ***Figures***:
  - o All figures included in the thesis named after the number used in the thesis.
- ***Tables***:
  - o All tables included in the thesis named after the number used in the thesis.

Additional Material:

- ***Figures:***
  - o ***Avalanche_data_set***
    Several plots processed of the data set of the reported avalanche data are included. The majority of plots is done for the data set including all avalanche records until the end of the season 2012/2013.
    - ▪ ***Distribution_avalanches***
      Each attribute is plotted in barplots. Each single value of the original data set as well as grouped values plotted in intervals are provided.
    - ▪ ***Distribution_attributes_per_region***
      Some attributes (Months, Release, Season, Weekdays, Years) are plotted in barplots, considering the amounts per SAIS area.
    - ▪ ***Frequencies_of_avalanche_days***
      Number of days with different numbers of avalanches are plotted for each area and in total.
    - ▪ ***Terrain_attributes_comparison_original_DEM***
      Comparisons of the original terrain attributes ‚Altitude', ‚Aspect' and ‚Gradient' of the original data set and those extracted from the DEM are providede.
    - ▪ ***Altitude_per_region***
      The altitude distribution in the bouding boxes of each SAIS area is plotted.
    - ▪ ***Relationships***
      Aspect is plotted against Altitude.
      Gradient is plotted against Altitude.
    - ▪ ***Terrain_attributes_and_avalanches***
      Distribution of the terrain attributes in the bounding boxes of the SAIS areas are plotted with the distribution of the avalanches within the attribute ranges.
    - ▪ ***Terrain_attributes_from_DEM***
      Distributions of the ‚Altitude', ‚Aspect' and ‚Gradient' values extracted of the DEM are plotted in barplots.
  - o ***Cluster_analysis***
    - ▪ ***Dendrograms***
      For each SAIS area, the dendrograms with thresholds for an avalanche day between >=1 and >=5 avalanches/day are plotted for the solutions with 4 clusters, 5 clusters and 6 clusters.
    - ▪ ***Distribution_of_resulting_variables***
      Distributions of several attributes of the total data and the data considering different thresholds for an avalanche days are plotted.
      Distributions of the variables of the scenario with a threshold >=3 avalanches/day are plotted for each SAIS area.
    - ▪ ***Figures_optimalScenarios_per_area***
      The figures included in the summary table of the optimal scenario for each area and the total amount of data (Boxplots, Maps, Temporal distribution, Terrain attributes) are provided.
  - o ***SnowProfile_data_sets***
    - ▪ ***Attributes_depth_information***
      The attributes included in the snow profile depth information are plotted in barplots.
      The attributes are plotted against the depth.
    - ▪ ***Attributes_resultingDataFile_fromPartB***
      Each attribute of the resulting data file of part B included in the header information is plotted in barplots for each SAIS area separately.
    - ▪ ***Comparison_data_sources***
      - • ***AirTemp_vs_SummitAirTemp***
        Air temperature values measured at the snow profile location are compared to the values measured at summit stations.
      - • ***Comparison_Group1_vs_Group2***
        Air temperature, wind direction and wind speed of different data sources are compared.
        Different scalings for the air temperature are plotted.
        Some further attributes (Precipitatin Code Id versus Snow Index, Snow Temperature, Maximum Hardness Difference and Maximum Temperature Gradient of Internet files and binary files) of different data sources are compared.
      - • ***Summit_data_sources_Cairngorms***
        Summit air temperature, summit wind direction and summit wind speed from different data sources are compared (Cairngorms station downloaded from the webpage [CGSummit], Cairngorm station data from the Floyer DVD [Floyer] and attributes included in the Snow Profile data set of the internet files [SP]).
    - ▪ ***Distribution_attributes_InternetFiles***
      Each attribute included in the snow profile internet data is plotted in barplots in theoriginal distribution as well as grouped in intervals.
    - ▪ ***Relationships_between_attributes***
      Relationships between some attributes (mainly air temperature, wind direction and wind speed measured at the snow profile locations and at summit stations) are compared in scatterplots.
    - ▪ ***SnowProfile_attributes_and_avalanches***
      Distribution of the total attributes and related number of avalanches that occurred at the related values are plotted.
    - ▪ ***Temporal_relationships_between_attributes***
      Summit air temperature and rain as well as snow index and rain are plotted against each other as seasonal time series for the seasons 2007/08 to 2012/13.
    - ▪ ***TimeSeries_2007_until_2013***
      Some attributes (Cloud, Drift, FootPen, MaxHardDiff, MaxTempGrad, PrecipCode, Rain, SnowIndex, SnowTemp, Temperature, WindDirection, WindSpeed) are plotted as seasonal time series for the seasons 2007/08 to 2012/13. The lines show the attributes per area, the points the number of avalanches and vertical gray lines indicate days on which at least one avalanche is reported.
- ***Resulting_data_sets***
  - o ***Combined_HeaderInformation_SnowProfiles.pdf***: Header Information of the snow profiles combined in part B
  - o ***Combined_DepthInformation_SnowProfiles.pdf***: Depth Information of the snow profiles combined in part B
  - o ***Metafile_HeaderAndDepthInformation_SnowProfiles.pdf***: Metafile describing the two resulting data sets in detail
  - o ***ReportedAvalanches_1990_2013.pdf***: Data set of the reported avalanches including the comment analysis and separate columns for the year/season/month/day/hour...

- WinPitReader
Java Program to read the Binary Files

Figure A.8: Overview of the content of the additional material (folder names are written in bold and italic).

A.8 R Code Examples

```
#-------------------------
# CLUSTERING
#-------------------------

#READ IN THE SNOW PROFILE DATA
Data <- read.csv(file="SP_Total_Kombi.csv", sep=";", header=TRUE)

#SELECT THE AVALANCHE DAYS USING THE THRESHOLD
Data.aval1 <- Data[with(Data,which(Data$NrOfAvalTotal >=1)),]

#SELECT ONLY CERTAIN VARIABLES
Data.use1 =
 Data.aval1[c("S_AirTemp_Kombi","S_AirTemp_3day","S_AirTemp_7day","S_WindSpeed_Kombi","S_Wind
 Speed_3day","S_WindSpeed_7day","SwIndex_neu","SwIndex_3day","SwIndex_7day","S_AirTemp_Diff",
 "Nr_Weakness")]

#ASSIGN UNIQUE ROWNAMES
rownames(Data.use1) <- paste(Data.aval1$Date, Data.aval1$Area, Data.aval1$ID,sep=" ")

#DELETE ROWS IN WHICH ONLY 1 TO 4 VALUES ARE INCLUDED
numNAs <- rowSums(is.na(Data.use1))
Data.use1 <- Data.use1[!(numNAs > 6),]

#CLUSTERING WITH STANDARDISATION AND WARD'S METHOD
Data.hclust1= hclust(dist(scale(Data.use1, center=TRUE, scale=TRUE)), method="ward")

#DEFINE THE NUMBER OF CLUSTERS
groups1 <- cutree(Data.hclust1,5)

#REORDER THE CLUSTERS
Data.res1 <- data.frame(Data.use1, groups1=factor(groups1))
Data.res1$groups1 = factor(Data.res1$groups1, levels=c(2,5,1,3,4));
levels(Data.res1$groups1) = c(1,2,3,4,5)

#-------------------------
# ANALYSE THE CLUSTERS
#-------------------------

#CALCULATE THE CLUSTER MEANS
clust.centroid = function(i, Data.use3, groups3) {
     ind = (groups3 == i)
     colMeans(Data.use3[ind,], na.rm=TRUE)
}
centroids <- sapply(unique(groups3), clust.centroid, Data.use3, groups3)

#SELECT ONLY DATA FROM ONE CLUSTER
Data1 <- Data.res1[which(groups1==1),]

#WRITE ROWNAMES OF DAYS IN ONE CLUSTER IN NEW VARIABLE
cluster <- names(subset(groups1, groups1==1))

#READ IN THE AVALANCHE DATA
Aval <- read.csv(file="Aval_Total.csv", sep=";", header=TRUE)

#SELECT ONLY AVALANCHES BELONGING TO A CERTAIN AREA AND A CERTAIN CLUSTER
Aval1 <- subset(Aval, Aval$Date %in% cluster & Aval$Region_new==3)

#SELECT ONLY AVALANCHES BELONGING TO A CERTAIN CLUSTER FOR THE TOTAL AMOUNT OF DATA
Aval1<- data.frame()
for(i in 1:length(cluster)) {
    Aval1<- rbind(Aval1, subset(Aval,grepl(strtrim(cluster[i],8),Aval$Date) &
    grepl(substring(cluster[i],10,10),Aval$Region_new)))
    }
```

Figure A.9: Example R code used for the clustering.

A.9 CA Plots

Figure A.10 includes the cluster sizes for each scenario shown in figures A.11 to A.34. Figures A.11 to A.34 provide all scenarios that are conducted for each SAIS area and the total amount of data for different thresholds for an avalanche day and for several numbers of clusters.

Northern Cairngorms: 4 Clusters

Threshold avalanche day	Cluster sizes
>= 1	68, **130**, 75, 130
>= 2	31, **42**, 30, 38
>= 3	6, 18, 5, **29**
>= 4	2, 5, 15, **13**
>= 5	5, 4, 3, **8**

Northern Cairngorms: 5 Clusters

Threshold avalanche day	Cluster sizes
>= 1	68, 39, 91, **130**, 75
>= 2	**31**, 31, 11, 38, 30
>= 3	6, **18**, 5, 7, 22 ✔
>= 4	2, 5, 6, **15**, 7
>= 5	**5**, 4, 3, 6, 2

Northern Cairngorms: 6 Clusters

Threshold avalanche day	Cluster sizes
>= 1	**68**, 39, 91, 73, 75, 57
>= 2	11, **31**, 11, 38, 30, 20
>= 3	6, 10, 5, 7, 8, **22**
>= 4	2, 5, 4, **11**, 7, 6
>= 5	1, **4**, 3, 6, 4, 2

Northern Cairngorms: 7 Clusters

Threshold avalanche day	Cluster sizes
>= 1	19, 39, 91, 57, 75, 73, 49
>= 2	11, 17, 11, 38, 30, 20, 14
>= 3	6, 10, 5, 7, 8, 10, 12
>= 4	2, 5, 4, 7, 7, 6, 4
>= 5	1, 3, 3, 6, 2, 4, 1

Lochaber: 4 Clusters

Threshold avalanche day	Cluster sizes
>= 1	**100**, 89, 143, 121
>= 2	49, **91**, 41, 52
>= 3	**35**, 17, 51, 22 ✔
>= 4	**27**, 10, 31, 8
>= 5	**18**, 4, 11, 6

Lochaber: 5 Clusters

Threshold avalanche day	Cluster sizes
>= 1	75, 25, 143, **121**, 89
>= 2	**49**, 33, 41, 58, 52
>= 3	20, 17, 51, **22**, 15
>= 4	14, 10, **31**, 8, 13
>= 5	6, 4, 12, 6, **11**

Lochaber: 6 Clusters

Threshold avalanche day	Cluster sizes
>= 1	75, 25, 143, 89, 61, 60
>= 2	21, 33, 41, 52, 58, 28
>= 3	20, 17, 51, 15, 12, 10
>= 4	14, 10, 26, 13, 8, 5
>= 5	6, 4, 12, 5, 6, 6

Glencoe: 4 Clusters

Threshold avalanche day	Cluster sizes
>= 1	**96**, 32, 66, 69
>= 2	**29**, 10, 36, 14
>= 3	12, 11, **13**, 11 ✔
>= 4	**5**, 4, 5, 11
>= 5	**3**, 4, 3, 4

Glencoe: 5 Clusters

Threshold avalanche day	Cluster sizes
>= 1	37, 32, 66, 69, **59**
>= 2	13, 16, **36**, 14, 10
>= 3	12, 11, 8, **11**, 5
>= 4	2, 4, 5, **11**, 3
>= 5	4, 2, **3**, 4, 1

Glencoe: 6 Clusters

Threshold avalanche day	Cluster sizes
>= 1	37, 32, 66, 69, 23, 36
>= 2	13, 16, 16, 14, 20, 10
>= 3	12, 11, 8, 7, 5, 4
>= 4	2, 4, 5, 9, 3, 2
>= 5	2, 4, 1, 4, 1, 2

Southern Cairngorms: 4 Clusters

Threshold avalanche day	Cluster sizes
>= 1	33, **30**, 25, 36
>= 2	**10**, 10, 12, 6
>= 3	2, **3**, 1, 5

Southern Cairngorms: 5 Clusters

Threshold avalanche day	Cluster sizes
>= 1	33, 25, 18, 36, 12
>= 2	6, 10, 12, 4, 6 ✔
>= 3	2, 2, 1, 1, 5

Creag Meagaidh: 4 Clusters

Threshold avalanche day	Cluster sizes
>= 1	93, **54**, 69, 83
>= 2	30, 23, 34, **65**
>= 3	24, 24, 14, **25**
>= 4	16, **22**, 16, 4
>= 5	9, 16, **9**, 3

Creag Meagaidh: 5 Clusters

Threshold avalanche day	Cluster sizes
>= 1	93, 28, 83, **69**, 26
>= 2	**30**, 23, 22, 34, 43
>= 3	**24**, 7, 18, 14, 24
>= 4	16, 4, 13, **16**, 9
>= 5	9, 3, **16**, 4, 5

Creag Meagaidh: 6 Clusters

Threshold avalanche day	Cluster sizes
>= 1	**93**, 28, 83, 29, 40, 26
>= 2	16, 23, 22, 34, **43**, 14
>= 3	15, 7, 18, 14, **24**, 9 ✔
>= 4	16, 4, **13**, 6, 10, 9
>= 5	**9**, 3, 9, 4, 5, 7

Creag Meagaidh: 7 Clusters

Threshold avalanche day	Cluster sizes
>= 1	20, 28, 83, 29, 40, 26, 73
>= 2	16, 23, 22, 34, 21, 14, 22
>= 3	15, 7, 18, 14, 9, 9, 15
>= 4	16, 4, 7, 6, 10, 9, 6
>= 5	5, 3, 9, 4, 5, 7, 4

Total: 5 Clusters

Threshold avalanche day	Cluster sizes
>= 1	**953**, 386, 910, 506, 1269
>= 2	**528**, 251, 678, 614, 401 ✔
>= 3	250, 297, **611**, 270, 341
>= 4	204, 138, 421, **280**, 200
>= 5	80, 164, **214**, 184, 229
>= 6	72, 59, 164, 81, 235
>= 7	20, 29, 128, 62, 163
>= 8	79, 32, 103, 70, 34
>= 9	60, 31, 52, 36, 59
>= 10	32, 14, 54, 25, 44
>= 11	21, 47, 12, 18, 18
>= 12	13,16, 16, 19, 9
>= 13	10, 8, 11, 19, 17
>= 14	11, 12, 10, 15, 7
>= 15	4, 8, 5, 16, 4

Total: 6 Clusters

Threshold avalanche day	Cluster sizes
>= 1	493, 386, 910, 506, **1269**, 460
>= 2	344, 251, **678**, 614, 401, 184
>= 3	250, 297, **339,** 270, 272, 341
>= 4	204, 138, **421**, 182, 98, 200
>= 5	80, **164**, 136, 184, 229, 78

Total: 7 Clusters

Threshold avalanche day	Cluster sizes
>= 1	493, 386, **910**, 506, 840, 460, 429
>= 2	344, 251, 255, **614**, 401, 184, 423
>= 3	250, **297**, 220, 270, 272, 341, 119
>= 4	**204**, 138, 272, 182, 98, 200, 149
>= 5	80, 113, 136, **184**, 229, 78, 51

Total: 8 Clusters

Threshold avalanche day	Cluster sizes
>= 1	493, 386, 326, 506, 840, 460, 429, 584
>= 2	344, 251, 255, 170, 401, 184, 423, 444
>= 3	250, 151, 220, 270, 272, 341, 119,146
>= 4	69, 138, 272, 182, 200, 135, 149, 98
>= 5	80, 113, 136, 58, 229, 78, 51, 126

✔	more detailed described scenario in tables A.11 to A.34
XXX	cluster that is divided when increasing cluster size

Figure A.10: Overview of the cluster sizes for the scenarios presented in figures A.11 to A.34

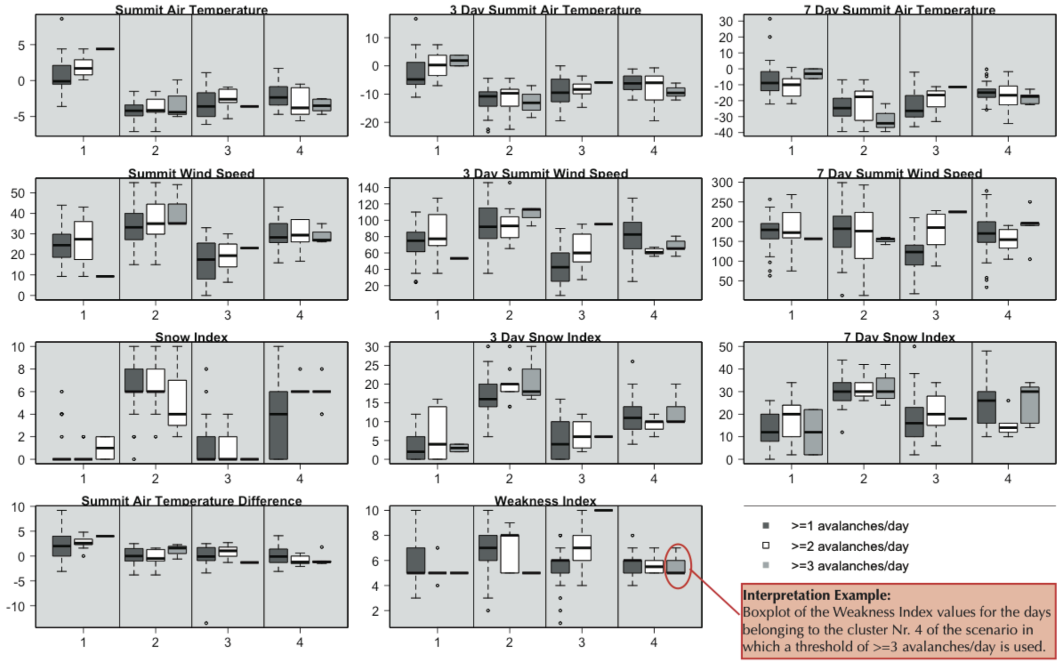


Figure A.11: Clustering scenarios with 4 clusters for the Southern Cairngorms.

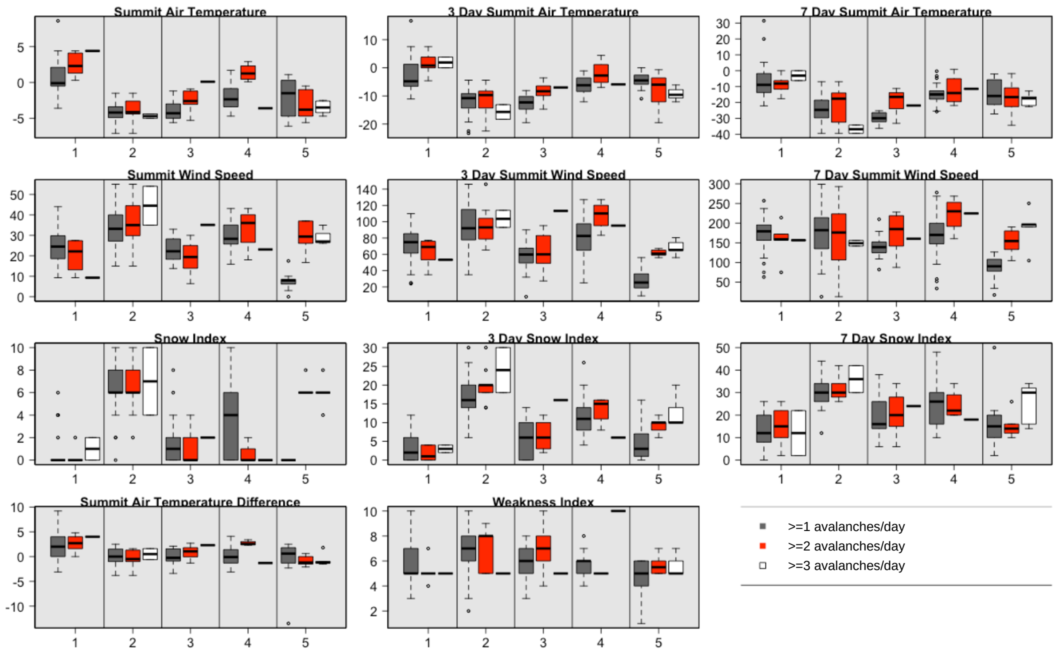


Figure A.12: Clustering scenarios with 5 clusters for the Southern Cairngorms.

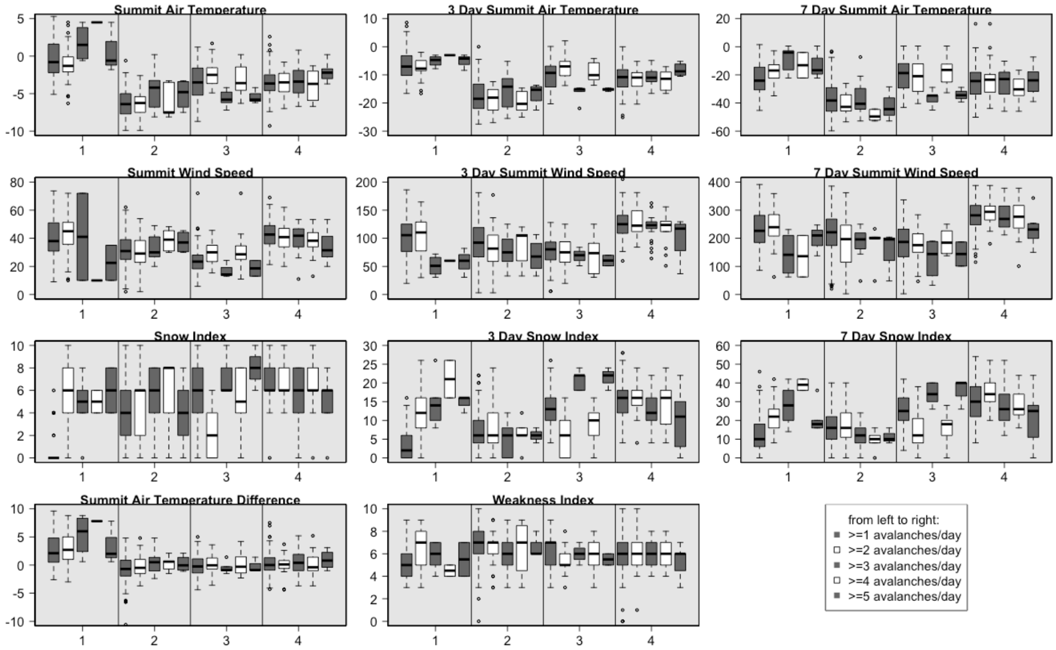


Figure A.13: Clustering scenarios with 4 clusters for the Northern Cairngorms.

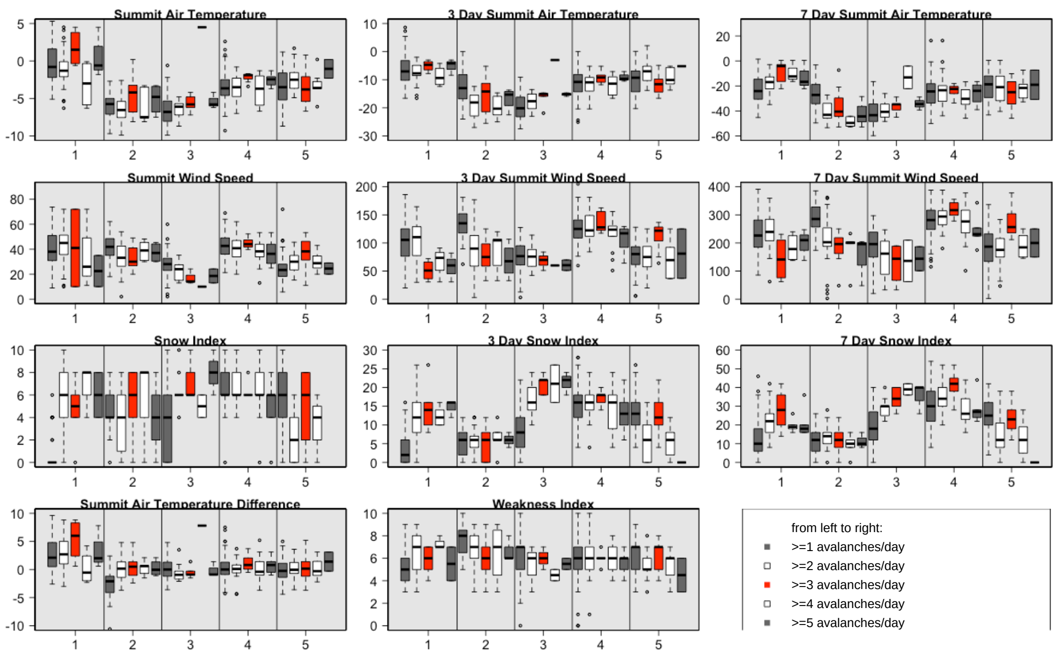


Figure A.14: Clustering scenarios with 5 clusters for the Northern Cairngorms.

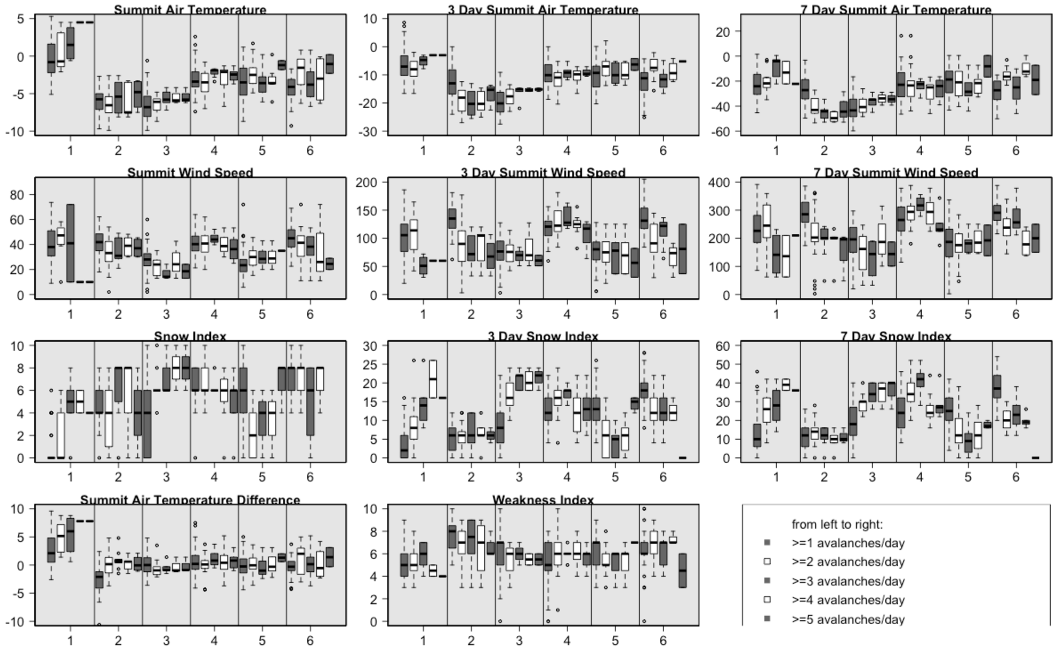


Figure A.15: Clustering scenarios with 6 clusters for the Northern Cairngorms.

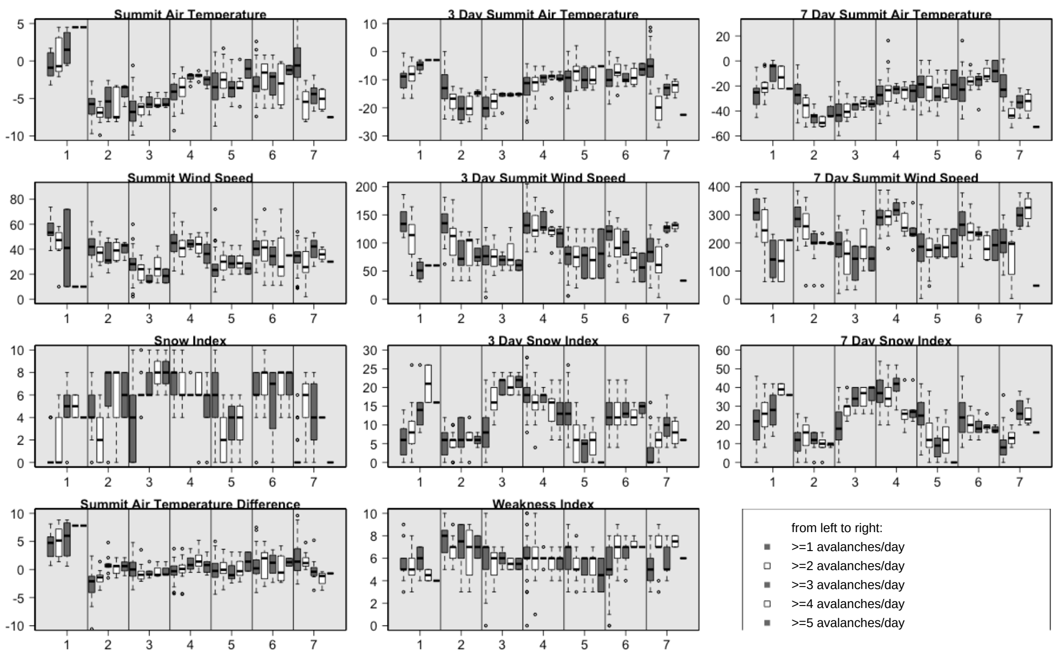


Figure A.16: Clustering scenarios with 7 clusters for the Northern Cairngorms.

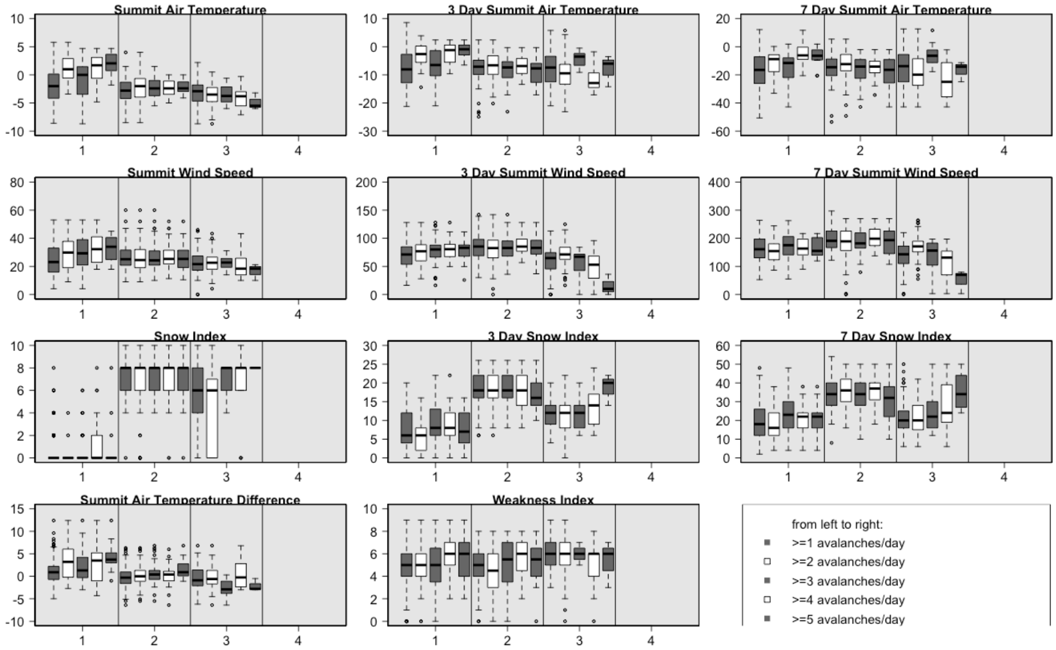


Figure A.17: Clustering scenarios with 3 clusters for Lochaber.

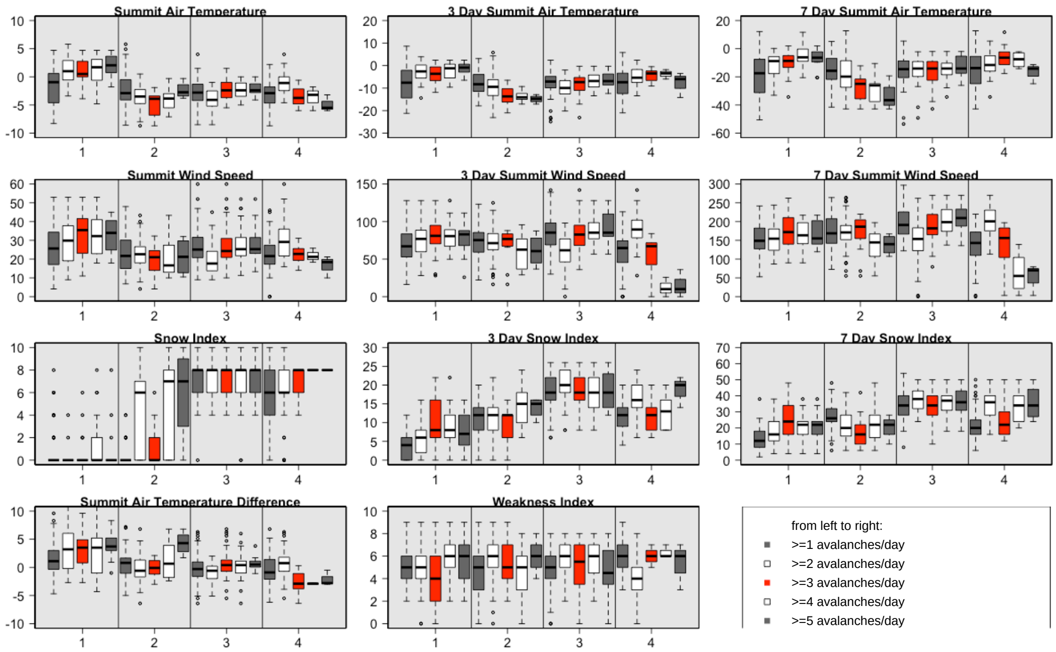


Figure A.18: Clustering scenarios with 4 clusters for Lochaber.

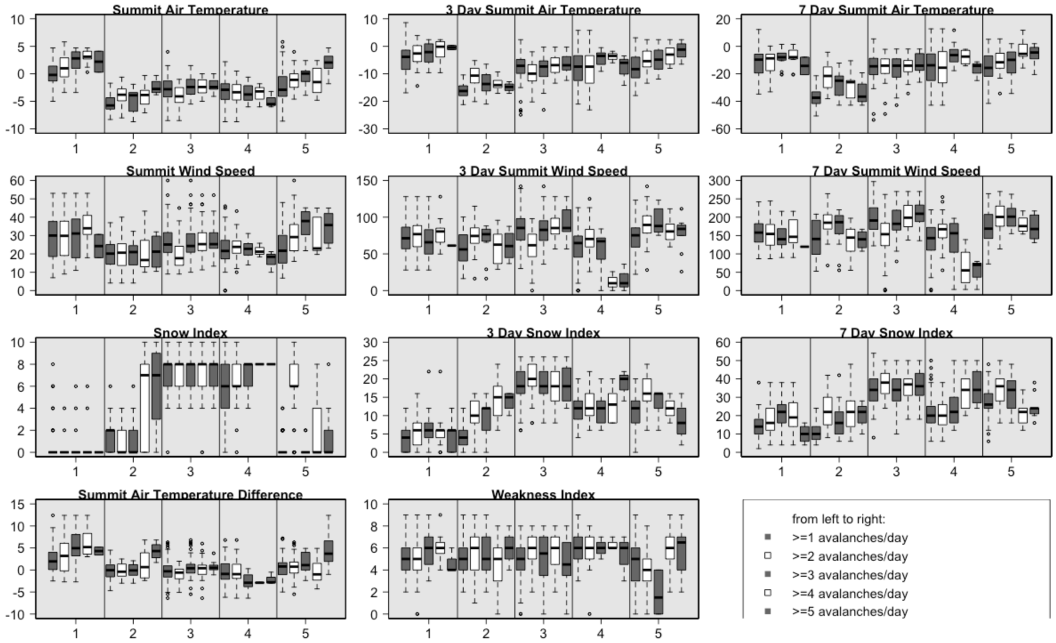


Figure A.19: Clustering scenarios with 5 clusters for Lochaber.

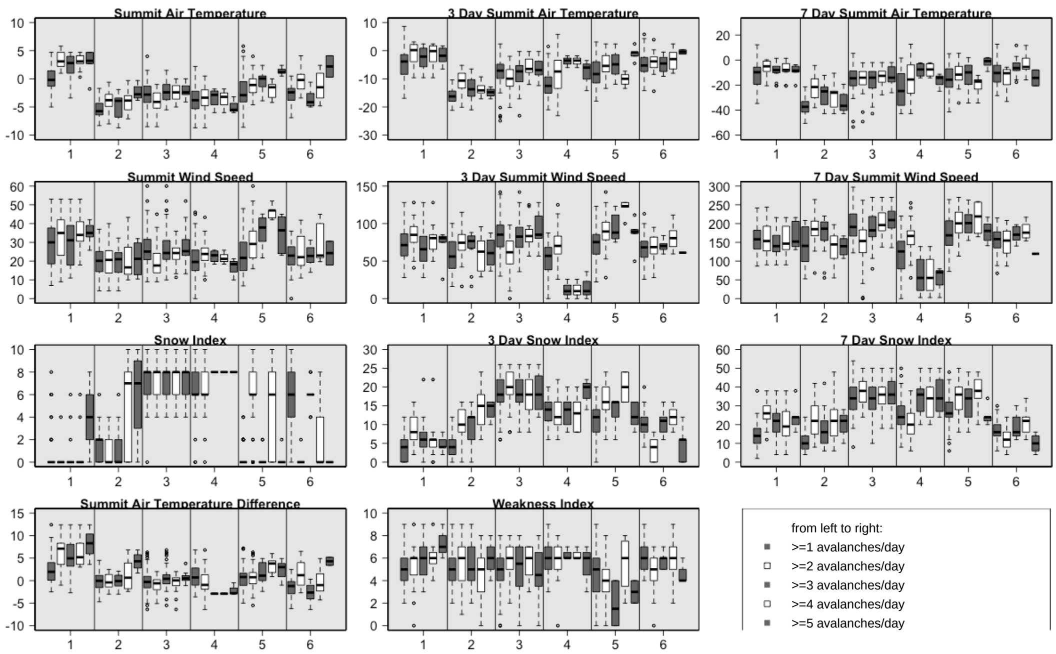


Figure A.20: Clustering scenarios with 6 clusters for Lochaber.

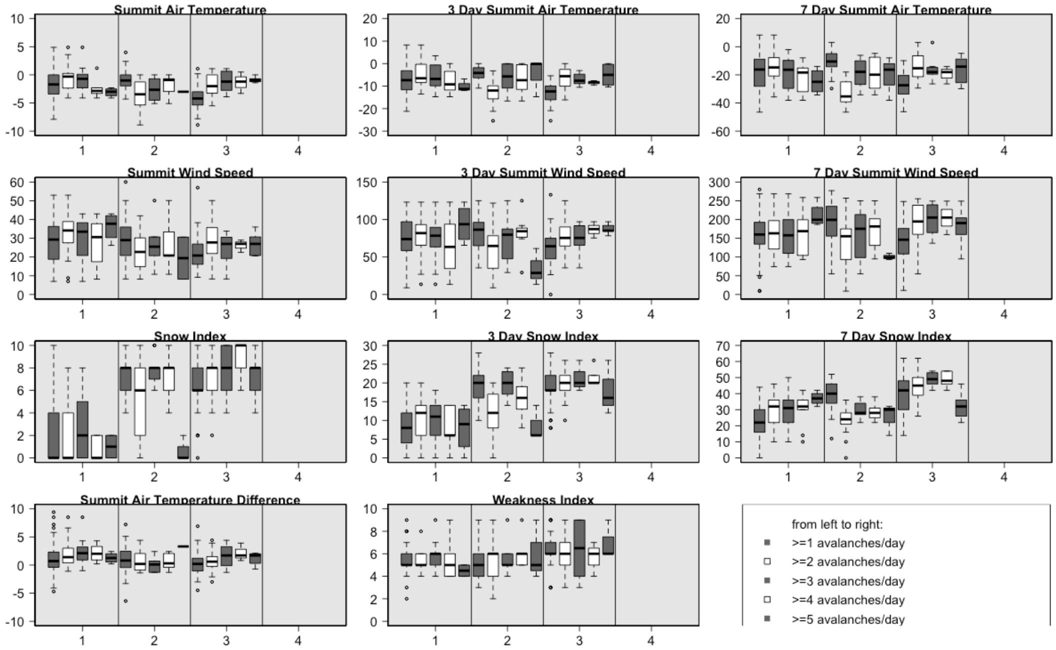


Figure A.21: Clustering scenarios with 3 clusters for Glencoe.

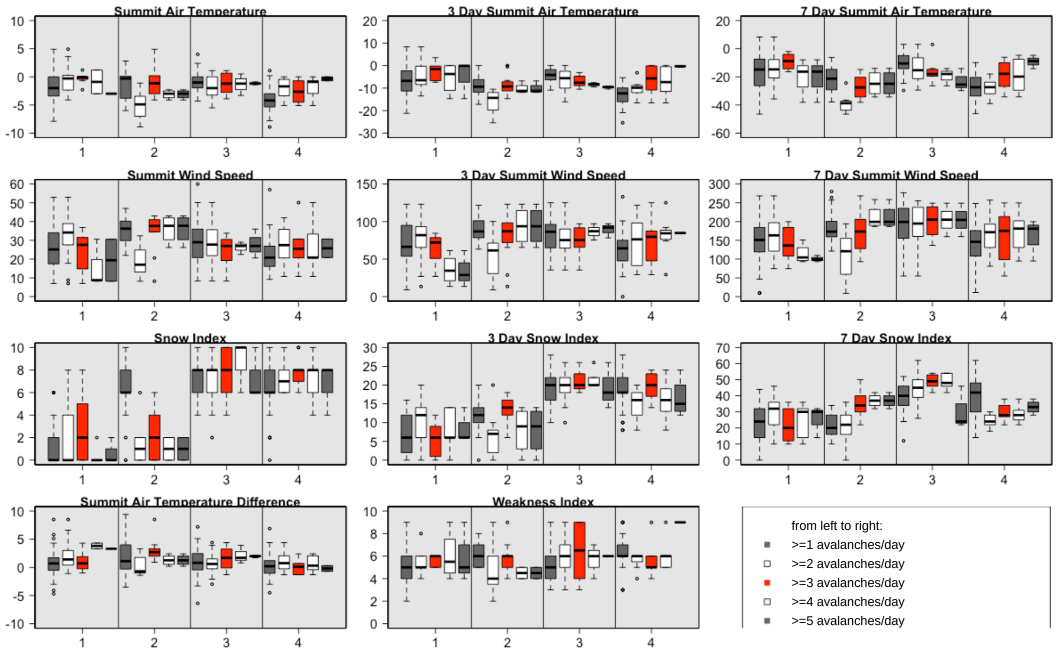


Figure A.22: Clustering scenarios with 4 clusters for Glencoe.

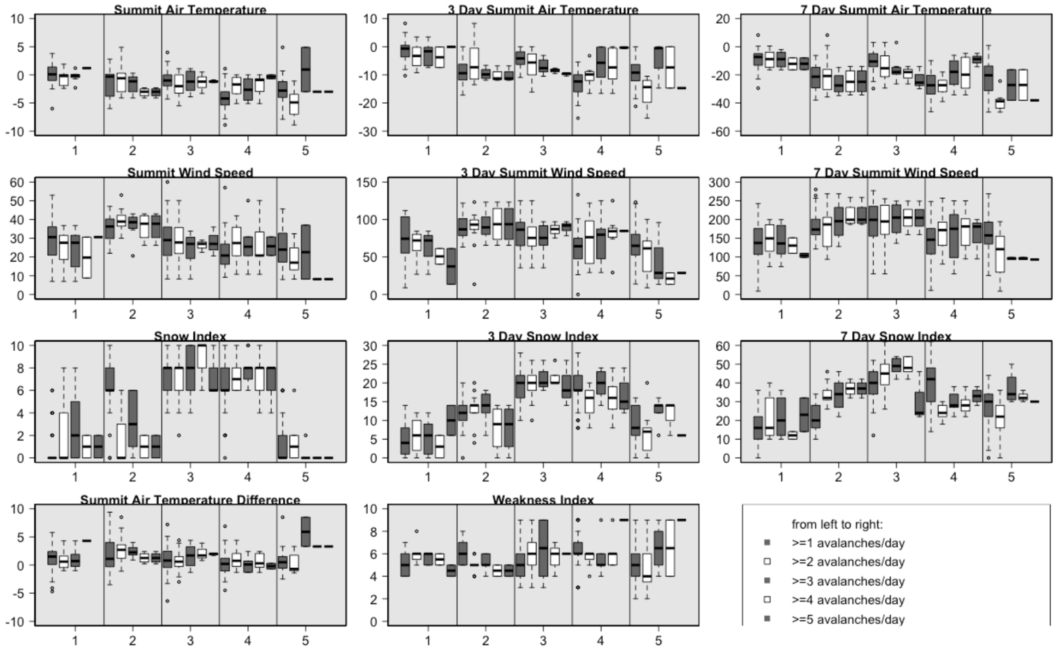


Figure A.23: Clustering scenarios with 5 clusters for Glencoe.

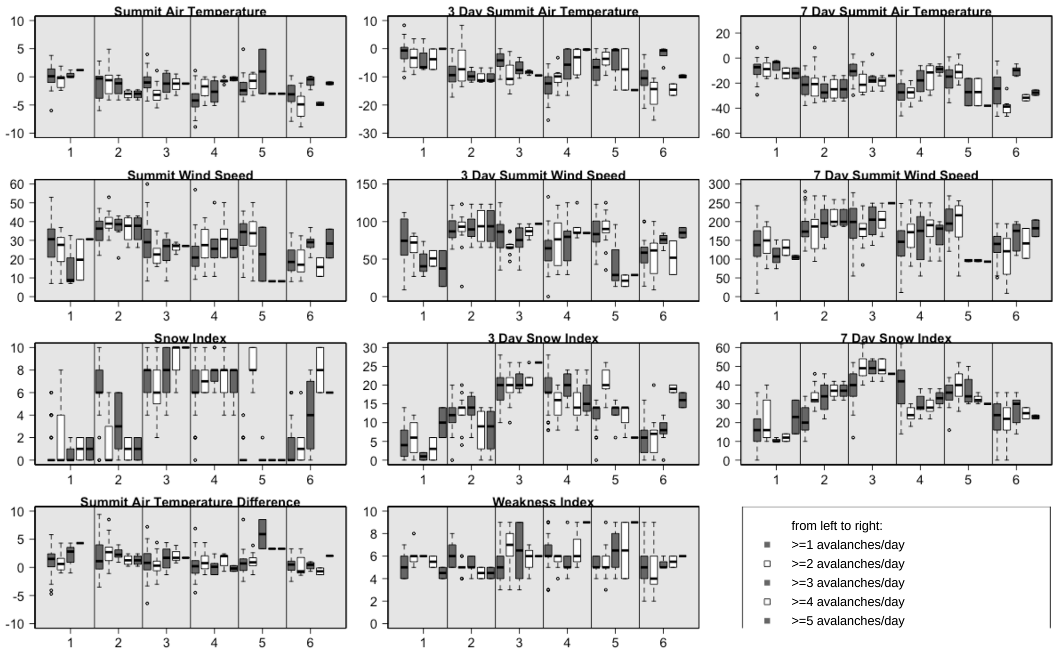


Figure A.24: Clustering scenarios with 6 clusters for Glencoe.

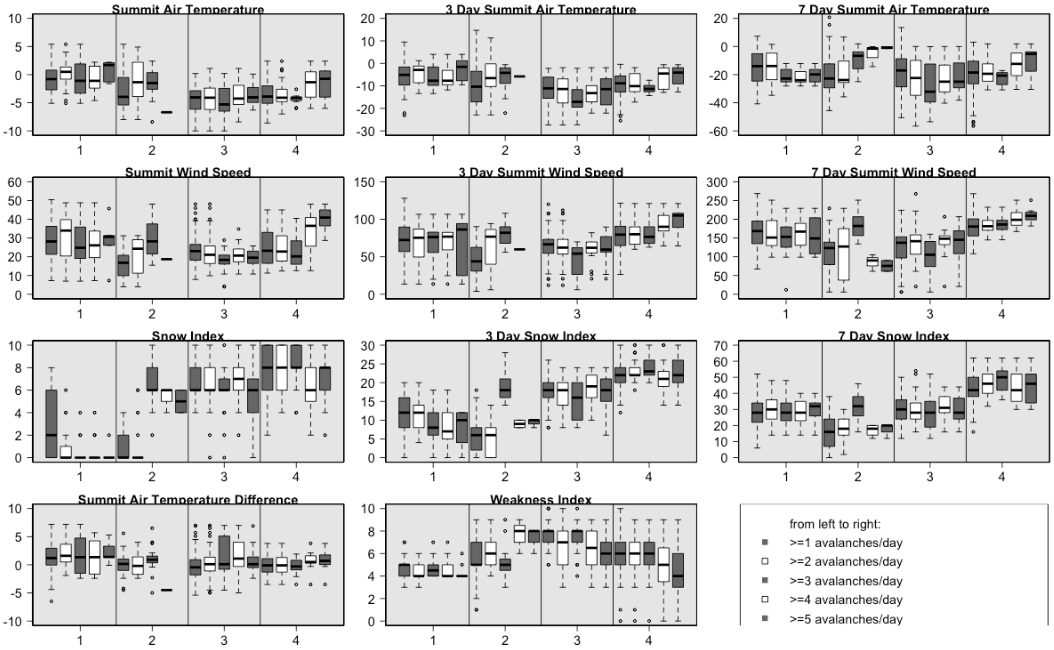


Figure A.25: Clustering scenarios with 4 clusters for Creag Meagaidh.

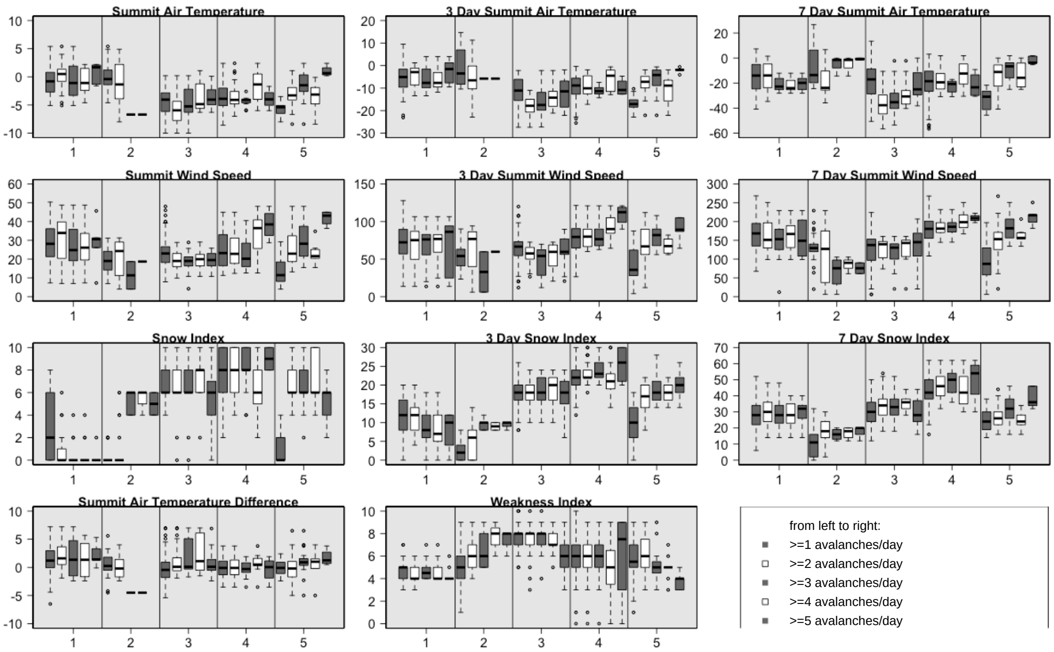


Figure A.26: Clustering scenarios with 5 clusters for Creag Meagaidh.

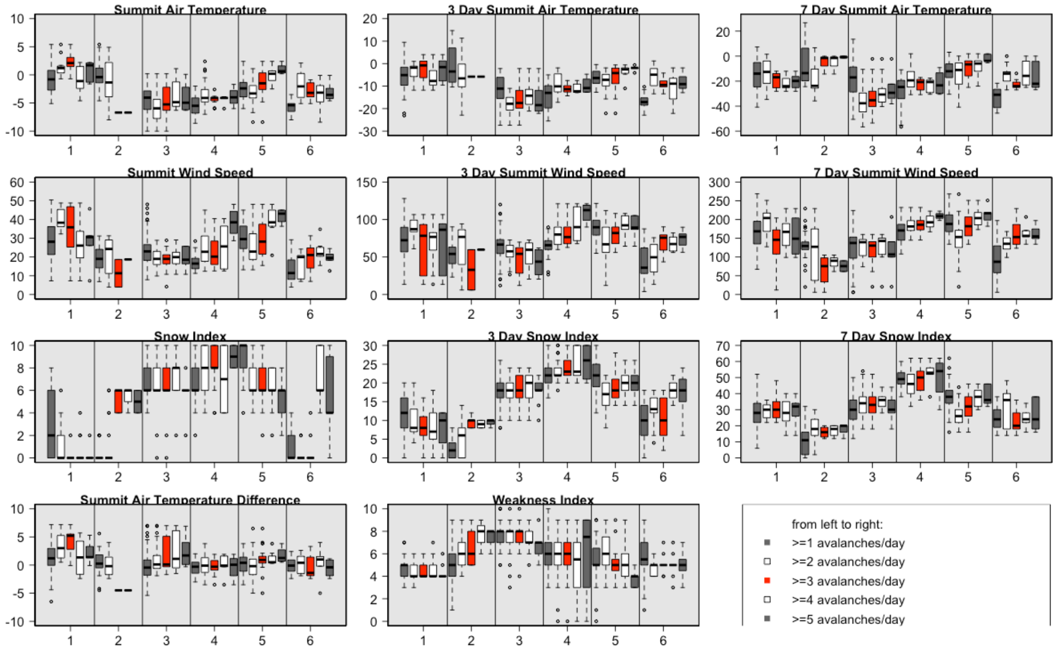


Figure A.27: Clustering scenarios with 6 clusters for Creag Meagaidh.

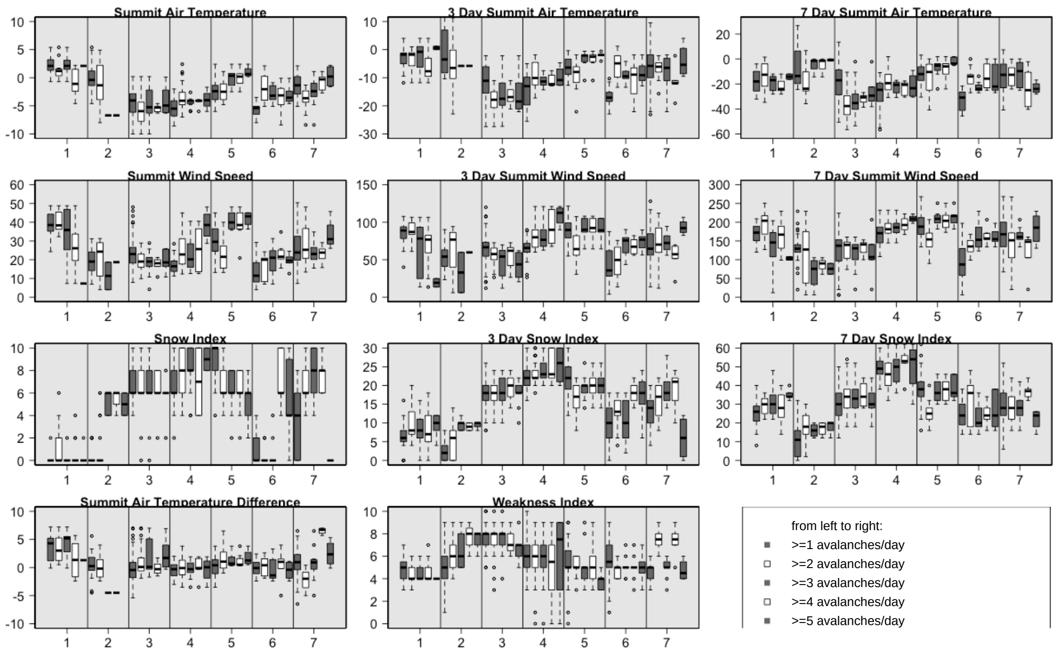


Figure A.28: Clustering scenarios with 7 clusters for Creag Meagaidh.

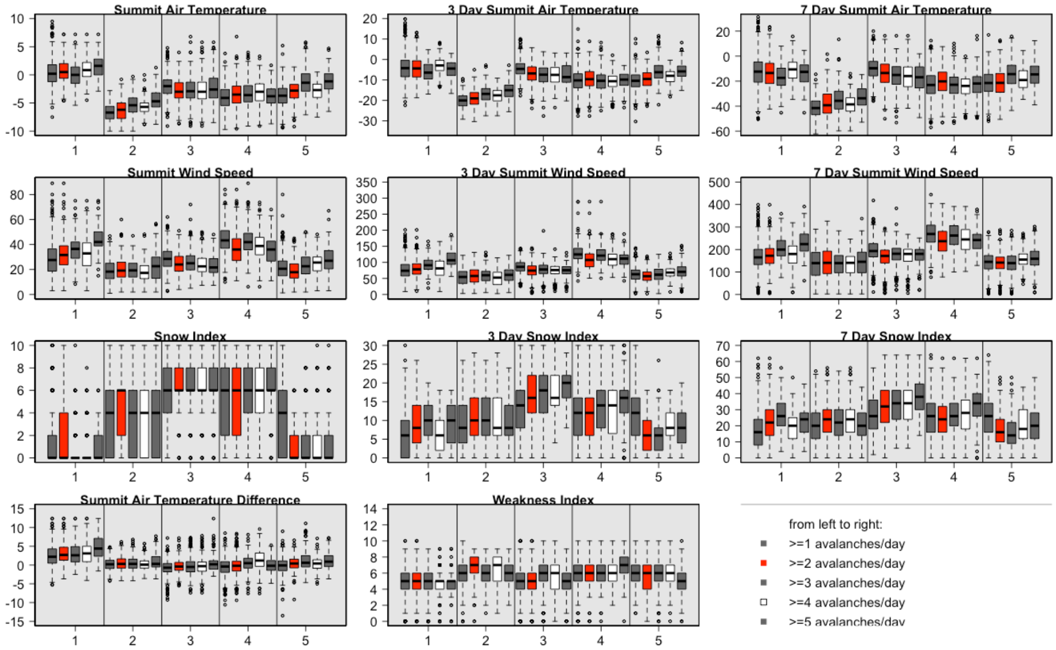


Figure A.29: Clustering scenarios with 5 clusters for the total amount of data.

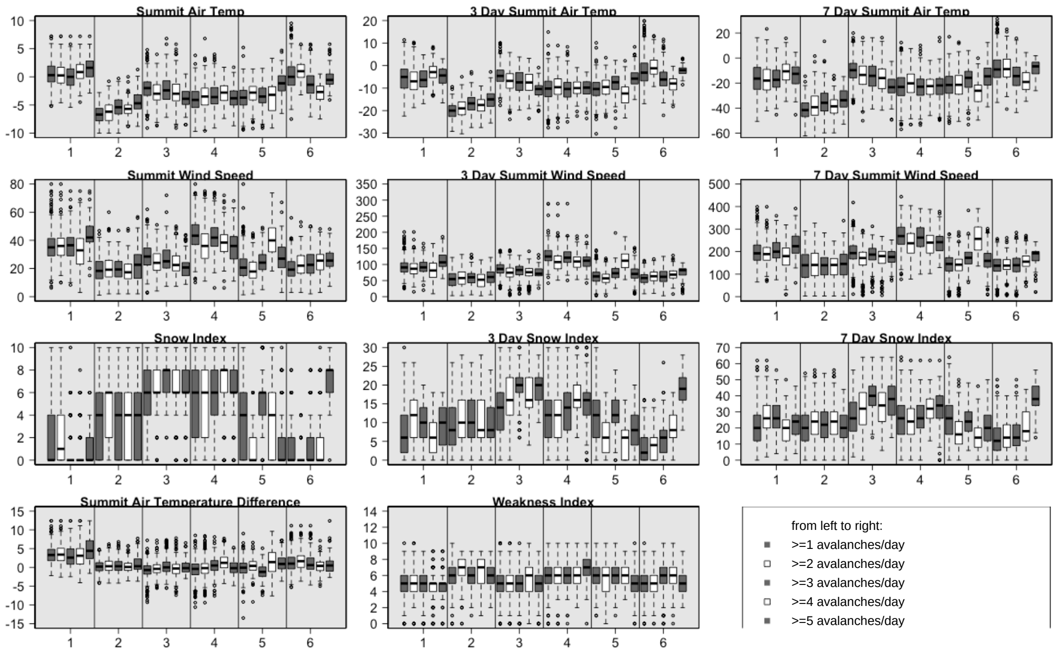


Figure A.30: Clustering scenarios with 6 clusters for the total amount of data.

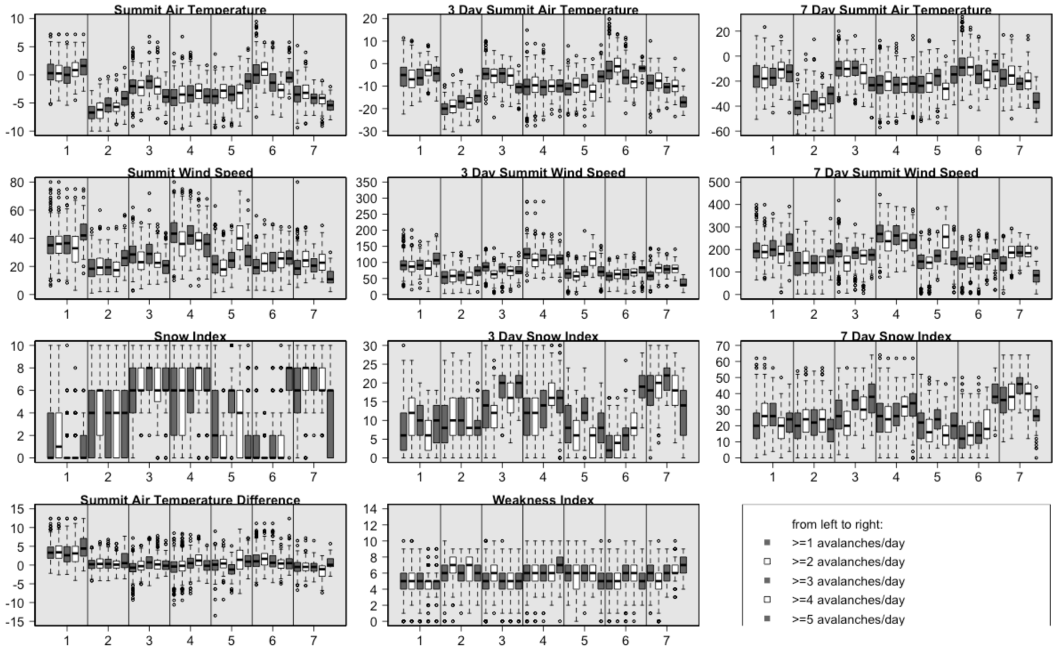


Figure A.31: Clustering scenarios with 7 clusters for the total amount of data.

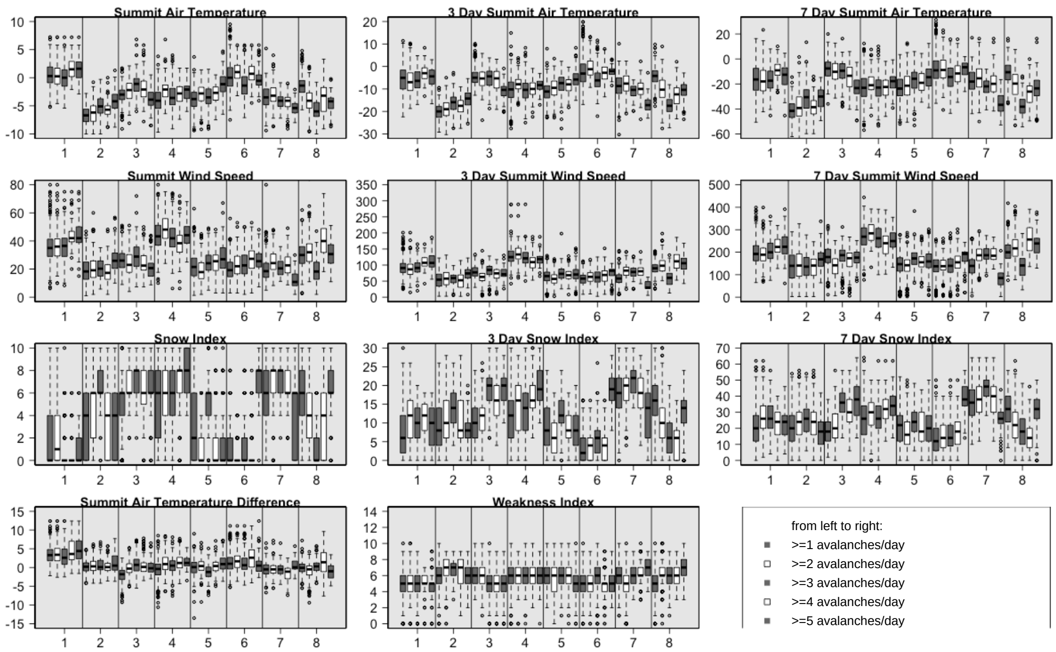


Figure A.32: Clustering scenarios with 8 clusters for the total amount of data.

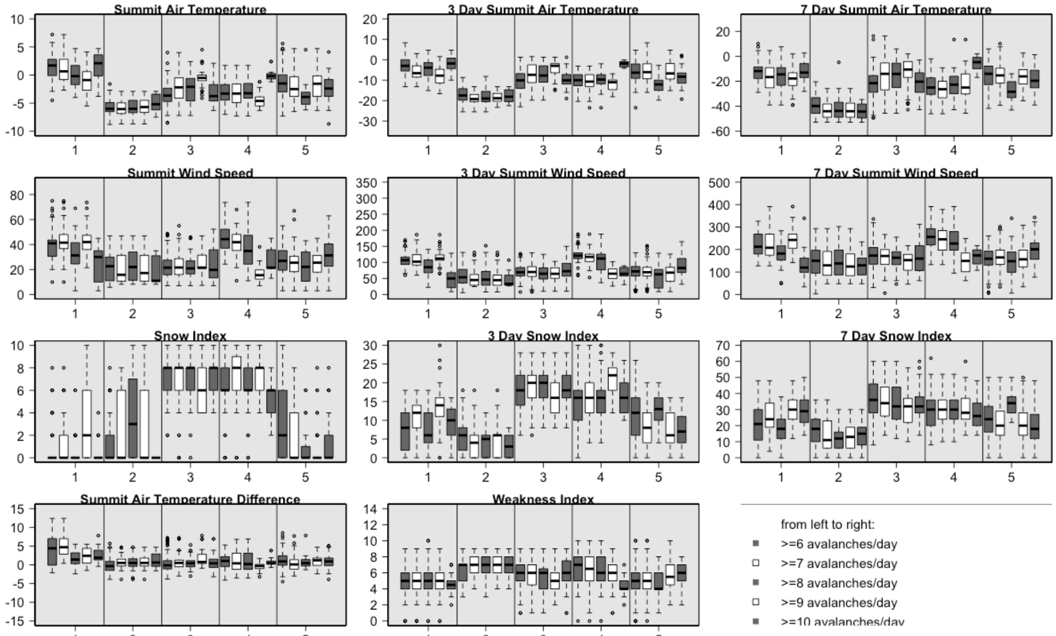


Figure A.33: Clustering scenarios with 5 clusters for the total amount of data using thresholds for an avalanche day between 6 and 10.

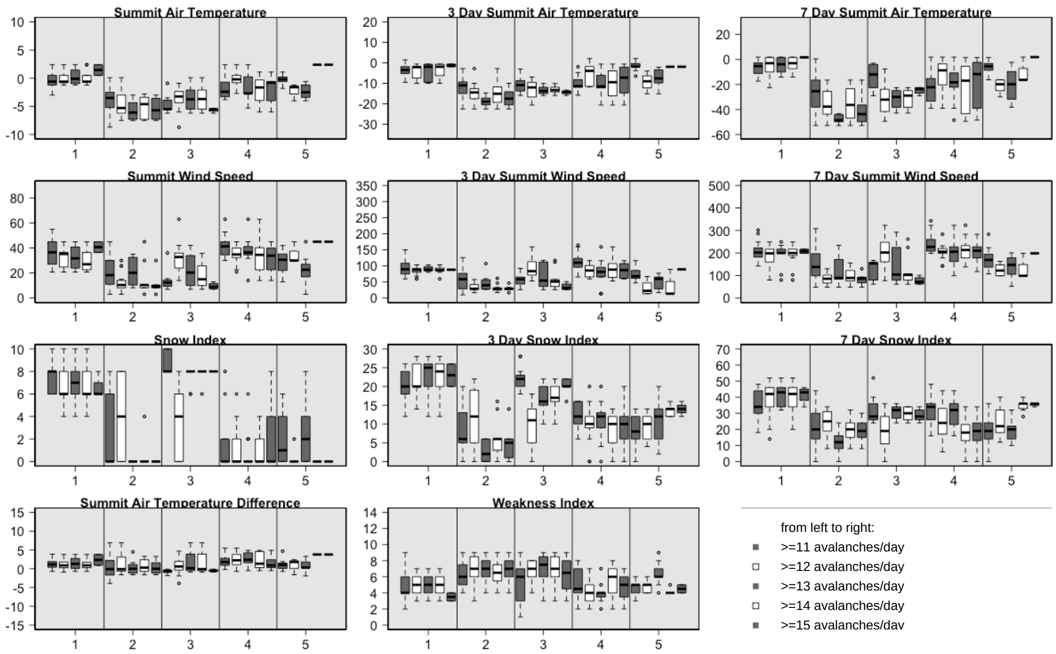


Figure A.34: Clustering scenarios with 5 clusters for the total amount of data using thresholds for an avalanche day between 11 and 15.

Personal Declaration

I hereby declare that the submitted thesis is the result of my own, independent, work. All external sources are explicitely acknowledged in the thesis.

Hinteregg, 27 September 2013