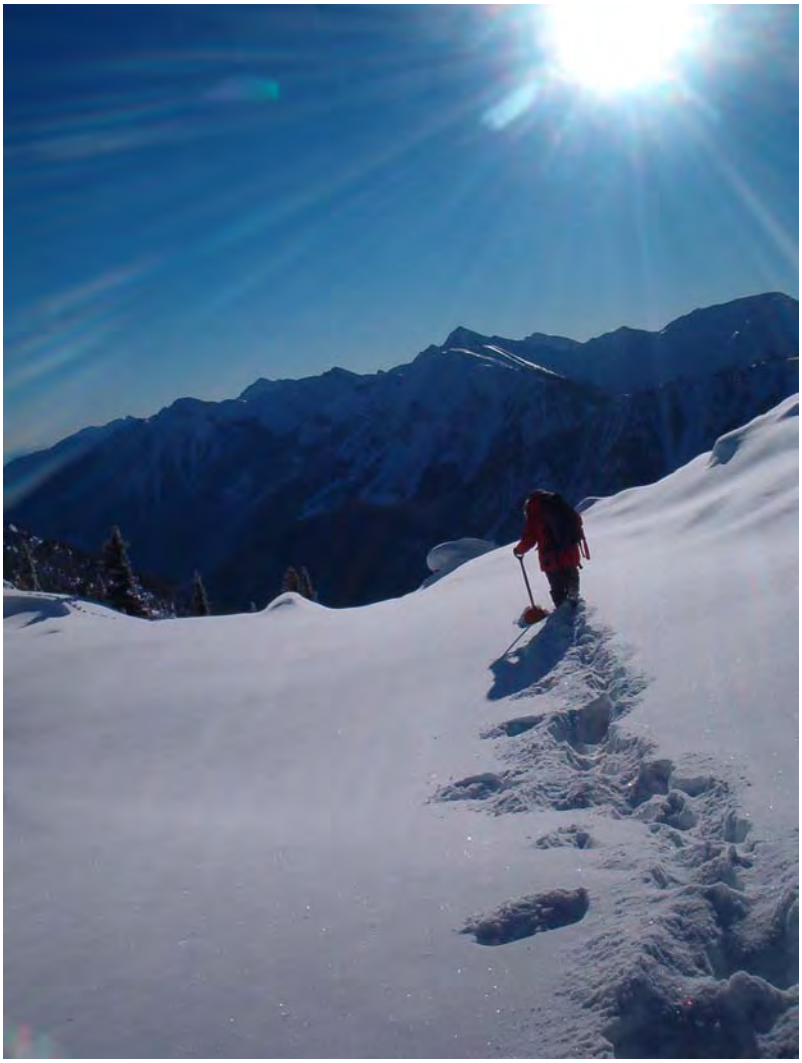

Forecasting Skier-Triggered Avalanches in the Columbia Mountains of Canada

Antonia Zeidler

UNIVERSITY OF CALGARY

Forecasting Skier-Triggered Avalanches in the Columbia Mountains of Canada

by

Antonia Zeidler

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CIVIL ENGINEERING

CALGARY, ALBERTA

November, 2004

Abstract

Computer assisted avalanche forecasting tools have proven to be valuable in some forecasting operations in Canada, mostly to predict natural avalanches. However most of the fatal avalanches in Canada are skier-triggered. The focus of this thesis was to improve the forecast of skier-triggered dry slab avalanches by incorporating stability indices and snowpack properties in addition to weather variables into an avalanche forecasting model. Up to 12 winters of data from near Blue River, an area north of Glacier National Park and within Glacier National Park in the Columbia Mountains of western Canada was analyzed.

Computer models for avalanche forecasting have made little use of snowpack properties; however, slab thickness (*H*), slab load (Load) and a skier stability index (Sk38) have shown promise for regional avalanche forecasting in the Columbia Mountains. Because it is not practical to measure the shear strength term in Sk38 daily, a method was refined for estimating the shear strength on days without manual snowpack measurements by adjusting the shear strength of persistent snowpack layers for normal load.

The main goal was to improve the forecast for skier-triggered avalanches on persistent weak layers, however non-persistent weak layers were analysed to run a forecasting model on a daily basis.

A daily skier instability index (DSI) was developed as a response variable using skier-triggered avalanches and stability ratings at the end of the day. The predictive merit of predictor variables was assessed using rank correlations, classification trees and a Nearest Neighbours program. Sk38 and Load, but not as much *H*, showed high potential to forecast DSI on a regional scale for persistent weak layers. In contrary the precipitation over the past 24 hours dominated in importance for forecasting skier-triggered avalanches on non-persistent weak layers.

Additionally a stability index for natural avalanches (Sn38) was assessed for its suitability to forecast avalanche activity at the highway corridor in Glacier National Park. Stability indices based on shear frame measurements have been used for several years by the avalanche control section; however the value of incorporating Sn38 in a computer model could not fully be assessed at this stage of the research.

Acknowledgements

First of I all I would like to thank Bruce Jamieson for the great supervision of my thesis, expertise and funding.

For their field work and company in the field I would like to thank: Michelle Gagnon, Alec van Herwijnen, Kyle Stewart, Paul Langevin, Ryan Gallagher, Ken Matheson and Cam Campbell.

For the great working environment, daily advice and support I thank the Avalanche Control Section of Glacier National Park: Dave Skjonsberg, Bruce McMahon, Jeff Goodrich, Johan Schleiss, Steve Thomas, Dean Flick, Jim Phillips, Tom Chalmers, Eric Dafoe and John Kelly.

Further my thanks go to Ross Purves for letting me use Cornice during my research.

I also thank the guides at Mike Wiegele Helicopter skiing for their feedback on my forecasting model.

For their field work in previous years I would like to thank: Jill Hughes, Leanne Allison, Tom Chalmers, Owen David, Mark Shubin, Ken Black, James Blench, Joe Filippone, Sue Gould, Aaron Cooperman, Brian Gould, Nick Irving, Crane Johnson, Greg Johnson, Alan Jones, Kalle Kronholm, Paul Langevin, Steve Lovenuik, Bill Mark, Greg McAuley, Rodden McGowan, Jordy Shepherd and Adrian Wilson.

For financial support we are grateful to the BC Helicopter and Snowcat Skiing Operators Association (BCHSSOA), the Natural Sciences and Engineering Research Council of Canada, Mike Wiegele Helicopter Skiing, Canada West Ski Areas Association (CWSAA) and the Canadian Avalanche Association. The supporting members of the BCHSSOA include Baldface Mountain Lodge, Bella Coola Heli Sports, Black Tusk Helicopter Inc., Canadian Mountain Holidays, Cariboo Snowcat Skiing and Tours, Cat Powder Skiing, Chatter Creek Mountain Lodges, Coast Range Helskiing, Crescent Spur Heli-skiing, Great Canadian Heli-skiing, Great Northern Snow Cat Skiing, Highland Powder Skiing, Island Lake Resort Group, Klondike Heliskiing, Last Frontier Heliskiing, Mica Heli Guides, Monashee Powder Adventures, Northern Escape Heli-skiing, Peace Reach Adventures, Powder Mountain Snowcats, Purcell Helicopter Skiing, R.K. Heli-

Skiing, Retallack Alpine Adventures, Robson Helimagic, Selkirk Tangiers Heli-Skiing, Selkirk Wilderness Skiing, Snowwater Heli-skiing, TLH Heliskiing, Valhalla Powdercats, Whistler Heli-Skiing and White Grizzly Adventures. The supporting members of Canada West Ski Areas Association include Apex Mountain Resort, Banff Mount Norquay, Big White Ski Resort, Hemlock Ski Resort, Intrawest Corporation, Kicking Horse Mountain Resort, Mt.Washington Alpine Resort, Silver Star Mountain Resorts, Ski Marmot Basin, Sun Peaks Resort, Sunshine Village, Whistler Blackcomb, Whitewater Ski Resort, and Resorts of the Canadian Rockies including Skiing Louise, Nakiska, Kimberley Alpine Resort, Fortress Mountain and Fernie Alpine Resort.

Thanks for last minute proof-reading to James Floyer, Dave Gauthier, Laura Bakersman, Alec van Herwijnen and Cam Campbell.

Special thanks go to Africa to Juliane Zeidler and Sharon Montgomery who proof-read, edited and encouraged me when time was running out.

I am grateful for my scholarship from the Gottlieb Daimler und Carl Benz Stiftung, Germany.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FREQUENTLY USED ABBREVIATIONS AND SYMBOLS

Generally used abbreviations

AD	CMH Adamants
ASARC	Applied Snow and Avalanche Research Group, University of Calgary
ASARC RP	ASARC field station at Rogers Pass in Glacier National Park
B	Bias
BC	British Columbia
BB	Bobbie Burns study plot
CAA	Canadian Avalanche Association
CMH	Canadian Mountain Holidays
CSI	Critical success index
CRUST	facets on or below crusts in the upper snowpack
CV	Coefficient of Variation
DSI	Daily skier instability index (DSI)
FAR	False-alarm rate
FC	Facets
GP	Glacier National Park
H	Hit rate
H	Slab thickness
ICSSG	International classification of seasonal snow on the ground
ISSW	International Snow Science Workshop
kPA	Kilo Pascal
KSS	Kuipers skill score
Load	Slab load
$Load_i$	slab load on the day to be forecast
$Load_{i-1}$	Slab load on the previous day
MW	Mike Wiegele Helicopter Skiing
NN	Nearest Neighbour model
NOV	November facets
NS	Near surface facets
p	statistical significance value
P_k	Ski penetration during skiing
PcpYLoad	Slab load in kPa over the last 24 hours as measured with a precipitation gauge
POD	Probability of detection
r^2	Coefficient of determination
RB	Rutschblock scores
S	Stability index for natural slab releases
S’	Equation for skier-triggered avalanches
SF	Stability factor
SH	Surface hoar
S_k	Modified skier stability index
Sn38	Extrapolated stability index for natural avalanches

TS	Threat score
UAA	Unweighted average accuracy

Chapter 4: abbreviations and conventions

C	Smooth, planar fracture surface
SBD	Small back divot (< 5 mm in depth)
MBD	Divot under rear compartment, 5-10 mm deep
BBD	Big back divot (> 10 mm in depth)
BC	Back divot extends beyond rear compartment
W	Wavy, more than one wave corresponding to frame sections (5-10 mm)
SW	Slightly wavy (wave height < 5 mm)
LC	Lateral chunk – fractured deeper at left or right side
IRR	Irregularities 5-10 mm deep
SIR	Irregularities < 5 mm deep
SH	Small hump (< 5 mm in height)
H	Hump (> 5 mm). Reject test, if frame pops up noticeably
STP	Fracture steps between two planes < 5 mm apart

Chapter 5: abbreviations and conventions

$\Sigma_{ø}$	Strength adjusted to the normal load
$ø(\Sigma, \sigma_{zz})$	Normal load adjustment
z	Distance from the snow surface normal to the slope
σ_{zz}	Normal load
ρ	Average slab density
g	Acceleration due to gravity
h	Slab thickness measured vertically
Ψ	Slope inclination, which is 0° in all the experiments used in this study
Pcp_{ij}	Daily loading rates
Σ_i*	A function of snowpack observations on day *i*;
$(\Delta\Sigma/\Delta t)_{ij}$*	A function of snowpack observations on day *i*
Σ_i*	Estimated shear strength on day *i* (kPa)
$\Delta t_{ij} = t_j - t_i$	Time interval between day *i* and day *j*
$(\Delta\Sigma/\Delta t)_{ij}$*	Estimated rate of change in shear strength (kPa d^{-1}) between day *i* and day *j*
Σ_j*	Forecast shear strength on day *j* (kPa)
rate $(\Delta\Sigma/\Delta t)_{ij}$	Shear strength change

Chapter 6: abbreviations and conventions

Age	Age of the weak layer (days); number of days since buried
Emax	Maximum grain size of weak layer crystal (mm)
Emin	Minimum grain size of weak layer crystal (mm)
H	Thickness of overlying slab (cm)
HH	Hand hardness exponent (Geldsetzer and Jamieson, 2001)

HS	Height of snowpack (cm)
Load	Weight per unit area of overlying snow (kPa)
SlabDens	Density of overlying slab (kg m^{-3})
Slope	Inclination of slope (°)
T+5	Temperature 5 cm above the weak layer (°C)
T-5	Temperature 5 cm below the weak layer (°C)
Ta	Air temperature (°C)
Ta/HS	Average snowpack temperature gradient (°C/m)
TG	Temperature gradient measured over 10 cm across failure plane of weak layer (°C/m)
Thick	Thickness of the weak layer (cm)
Twl	Temperature of the weak layer (°C)

Chapter 7: abbreviations and conventions

Baro	Barometric pressure at a.m. (mbar)
dBaro	Change in barometric pressure in past three hours. 1 = steady, 2 = rising, 3 = falling, 4 = rising rapidly, 5 = falling rapidly (CAA, 2003: 10)
dT	Change in a.m. air temperature from previous day (°C)
H	Slab thickness above a weak layer (cm)
HNY	Height of new snowfall for previous 24 h (m)
HS	Height of snowpack at a.m. (m)
HSTD	Height of snow that has accumulated since the last scheduled observation (once or twice a week)
HSTM	Height of snow accumulated on a board since the beginning of a storm
Load	The weight of overlying slab above a weak layer (kPa)
NaPrev	Number of natural avalanches on previous day
Pcp	1 = light rain, 2 = nil, 3 = snow < 1 cm/h, 4 = snow ~ 1 cm/h, 5 = snow ~ 2 cm/h, 6 = snow ~ 3 cm/h (CAA, 2003: 3)
PcpY	Water equivalent of precipitation on previous day (mm)
RH	Relative humidity at a.m. (%)
RHmnY	Minimum relative humidity for previous day (%)
RHmxY	Maximum relative humidity for previous day (%)
Sk38	Skier stability index
Sky	1 = clear, 2 = scattered, 3 = broken, 4 = overcast, 5 = obscured (CAA, 2003: 2)
Strm	Cumulative new snowfall (storm) since last day with less than 0.3 mm of precipitation (m)
T	Air temperature at a.m. (°C)
TmaxY	Maximum temperature for previous day (°C)
TminY	Minimum temperature for previous day (°C)
TriPrev	Number of skier-triggered dry slab avalanches on previous day
WrunY	24 h wind run for previous day (km)
WS	Wind speed at a.m. (km h^{-1})
WSa	Average upper air wind speed (km h^{-1})

WSad	1 = calm, 2 = light, 3 = moderate, 4 = strong, 5 = extreme (CAA, 2003: 9)

Chapter 8: abbreviations and conventions

HNY	Height of new snowfall for previous 24 h (m)
HS	Height of snowpack at a.m. (m)
PcpY	Water equivalent of precipitation on previous day (mm)
RH	Relative humidity at a.m. (%)
RHmnY	Minimum relative humidity for previous day (%)
RHmxY	Maximum relative humidity for previous day (%)
Sk38	Skier stability index
Strm	Cumulative new snowfall (storm) since last day with less than 0.3 mm of precipitation (m)
Strm3day	Cumulative new snowfall (storm) over the last three days (m)
T	Air temperature at a.m. (°C)
TmaxY	Maximum temperature for previous day (°C)
TminY	Minimum temperature for previous day (°C)
TriPrev	Number of skier-triggered dry slab avalanches on previous day
TriPrevBR	Skier-triggered avalanches in storm snow on previous day (yes/no)
WD	Wind direction at a.m.; (minus 90° modulo 360) east as base azimuth (°)
WDavg	Average wind direction over 24 hours; (minus 90° modulo 360) east as base azimuth
WrunY	24 h wind run for previous day (km)
WS	Wind speed at a.m. (km h^{-1})
WSavg	Average wind speed over 24 hours
WV	Wind vector, magnitude of the direction and intensity of the wind at a.m.
WVavg	Average wind vector over 24 hours

Chapter 9: abbreviations and conventions

HNY	Height of new snowfall for previous 24 h (m)
HS a.m.	Height of snowpack at a.m. (m)
NaPrev (AvalY)	Number of natural avalanches on the previous day
PcpY	Water equivalent of precipitation on previous day (mm)
RH a.m.	Relative humidity at a.m. (%)
RHmnY	Minimum relative humidity for previous day (%)
RHmxY	Maximum relative humidity for previous day (%)
Sn38	Stability index
Strm	Cumulative new snowfall (storm) since last day with less than 0.3 mm of precipitation (m)
Strm3day	Cumulative new snowfall (storm) over the last three days (m)
T a.m.	Air temperature at a.m. (°C)
TmaxY	Maximum temperature for previous day (°C)

TminY	Minimum temperature for previous day (°C)
WD a.m.	Wind direction at a.m.; (minus 90° modulo 360) east as base azimuth (°)
WDavg	Average wind direction over 24 hours; (minus 90° modulo 360) east as base azimuth
WrunY	24 h wind run for previous day (km)
WS a.m.	Wind speed at a.m. (km h^{-1})
WSavg	Average wind speed over 24 hours

1. Introduction

1.1 Development of the problem under investigation

Over the last century the number of avalanche victims on roads or in buildings has decreased significantly in Canada, whereas the number of avalanche deaths during recreational activities such as skiing, has risen. Regrettably this trend peaked in the winter of 2002/2003 when 29 people lost their lives in avalanches in the backcountry due to an exceptionally unstable snowpack. The general trend may be a result of the increasing number of people using the backcountry in winter while the decrease in the number of victims on roads and in buildings can be related to the establishment of forecasting operations and avalanche protection structures. Even though the number of avalanche victims during recreation has increased over the years, the trend is still low in proportion to the increase of people using the backcountry (Jamieson and Geldsetzer, 1996: 7). Good avalanche safety measures in commercial backcountry skiing operations, better avalanche education and better public warning services have contributed to backcountry users' avalanche awareness.

Studies on avalanche accidents show that the victim or group members often trigger the avalanche themselves (Jamieson and Geldsetzer, 1996: 10), which emphasizes the importance of forecasting skier-triggered avalanches. In the remainder of this thesis the expression skier-triggered refers to all avalanches triggered by skiers, snowboarders, hikers, etc. In most of the avalanche accidents the weak layers on which the avalanche is released can be identified and analysis shows that the weak layers exhibit specific characteristics during the course of the winter. This information can be used in daily stability evaluations.

Computer assisted avalanche forecasting tools have proved to be valuable in some forecasting operations in Canada. So far, computer programs are mainly used for data management and data visualization, whereas the use of data analysis and data exchange is presently limited. Most analysis tools still need to run more efficiently in operational mode and need to be configured for forecasting purposes in a specific forecasting area. Fortunately due to better computing capacities the models have become more and more flexible and adaptable to different forecasting areas. However, the models have to prove

that they can complement experience-based forecasting before operational use will spread. In addition, most forecasting models predict natural or a combination of natural and artificially triggered avalanches using mainly meteorological data, while Jamieson (1995: 147-155) has shown that a stability index based on manual snowpack measurements is the most significant indicator for skier-triggered avalanches on persistent weak layers on a regional scale.

Before formulating the objectives of this thesis (Section 1.7) an overview of the relevant snowpack processes, avalanche characteristics and avalanche protection measures will be given.

1.2 Snowpack processes

Once snow falls onto the ground it goes through significant changes to build up a variable snowpack. As a result of different weather events throughout the winter a layered snowpack (including weaker and stronger layers) forms (Colbeck, 1991). The layers differ in grain form, grain size, temperature, density, thickness, water content, hardness and therefore in strength. The snowpack properties change over time and consequently the snowpack may gain or lose strength as the winter progresses. Furthermore, the snowpack is spatially variable due to micro scale processes influenced, for example, by the aspect, inclination, ground cover and wind direction. It is important to understand snowpack properties and the processes that affect them, because they are the key to stability and avalanche forecasting.

Thermodynamic processes, and to a certain extent the overburden pressure, drive the metamorphism of the snow crystals in the seasonal snowpack (McClung and Schaerer, 1993: 48). The stored heat in the ground from summer and geothermal heat influence the temperature at the base of the snowpack whereas diurnal fluctuations and the prevailing synoptic conditions influence the temperature in the upper snowpack (McClung and Schaerer, 1993: 46). In general, water vapour moves from high to low vapour pressure areas (i.e. from warm to cold), which follows the assumption that the pore spaces are at or near saturation. In the snowpack this means a general heat flow from the lower to the upper snowpack because the temperature at the base of the snowpack is around 0°C due to the

heat flow from the ground, whereas the air temperature is below freezing for most of the winter. The temperature gradient (°C/m) within the snowpack is important for the metamorphism of the snow crystals and determines the existence of weak snowpack layers. It is only in spring that the snowpack may become isothermal, i.e. the entire snowpack temperature is at 0°C, and consequently no significant temperature gradient exists. In general, four types of metamorphism can be distinguished:

1) The *reduction of surface-to-volume ratio* generally affects newly fallen snow and is a result of the physical law that particles try to minimize their surface energy. The vapour pressure over convex surfaces is higher than over concave surfaces, so that the branches of new snow crystals round off (McClung and Schaerer 1993: 46). Furthermore, smaller particles tend to disappear due to a higher water vapour pressure whereas larger crystals grow. This type of metamorphism is considered an early state of rounding although the reduction of the surface-to-volume ratio may occur in colder temperatures than is common for rounding (see *equilibrium metamorphism*). This reduction of surface-to-volume ratio, caused by vapour pressure differences on the surface of a crystal, plays a minor role in forming the structure of the snowpack compared to the importance of temperature and temperature gradient driven forms of metamorphism.
2) During *equilibrium metamorphism* or *rounding*, the snow crystals reduce in maximum length and surface area per unit volume. Furthermore, the number of crystal contacts increases and the pore space between crystals decreases as the layer densifies (McClung and Schaerer, 1993: 54). Rounding is common when snowpack temperatures are at or below 0°C and is faster when temperatures are close to the melting point. At around 0°C melt-freeze metamorphism begins. In general equilibrium metamorphism stabilizes the snowpack.
3) Facets and depth hoar are the characteristic crystal forms that develop during *kinetic growth* or *faceting*. Usually these crystals have few contacts and smaller bonds between the grains, may grow large and have big pore spaces between them. This results in a less cohesive, often less stable snowpack. Kinetic growth

occurs at high temperature gradients of about 10°C/m or greater and is faster at relatively warmer temperatures, and more intensive where larger pore spaces allow undisturbed growth (McClung and Schaerer 1993: 49). Because this process is faster in warmer temperatures, depth hoar - an advanced form of faceting - typically develops near the ground where the temperature is close to melting point due to the heat flow from the ground. Because of its great importance for avalanche forecasting, and for this thesis, a detailed description of facet growth in the snowpack will be given in Section 1.4.2.

4) During *melt-freeze metamorphism* liquid water assembles in concave areas (lower pressure) and freezes to produce rounded melt forms. Again, smaller grains tend to disappear whereas larger crystals tend to grow due to the different melt temperatures. This process occurs at temperatures around 0°C. This type of metamorphism is most relevant in spring and will not be analyzed in this thesis.

In summary, the type of metamorphism and consequently the crystal form and size depends mainly on a) the temperature gradient, b) the temperature, and c) the pore space (McClung and Schaerer 1993: 49).

As mentioned above, the snowpack consists of a sequence of weaker and stronger layers and the weaker layers play a major role in avalanching as will be discussed in Section 1.4.

1.3 Types of avalanches

Avalanches are sudden snow movements on slopes. The two main types of avalanche are:

1) *Loose snow avalanches* or *point releases* (see Figure 1.1). These generally start at one point when the slope angle exceeds the critical static friction angle necessary to cause motion (defined for each type of snow), gather mass on the way down and result in a triangular shape (McClung and Schaerer, 1993: 72). Usually loose snow avalanches start at the surface of the snowpack or in near surface snow and involve low cohesion snow (McClung and Schaerer, 1993: 61). This type of avalanche is potentially less dangerous to skiers than slab avalanches because the snow is less

cohesive and the avalanches often smaller. Loose snow avalanches accounted for 5% of avalanche accidents investigated between 1984 and 1996 (Jamieson and Geldsetzer 1996: 16) but are not investigated in this thesis.

Figure 1.1: Loose snow avalanche (photo: B. Jamieson)

2) *Slab avalanches* have a distinct fracture line, flanks, a bed surface and a stauchwall (see Figure 1.2). The crown represents the top of the slab and is formed by a tension fracture, the bed surface is the sliding surface of the slab, the flanks represent the sides of the slab and the stauchwall is the lowest down-slope fracture surface (McClung and Schaerer, 1993: 75). The slab is a cohesive layer of snow and releases on a weaker failure layer beneath it (McClung and Schaerer 1993: 75). Ninety five per cent of all recreational accidents from 1984 to 1996 involved slab avalanches (Jamieson and Geldsetzer 1996: 16). Slab avalanches are potentially

more harmful to skiers, on the one hand because they are harder to predict than loose snow avalanches, which occur mostly during or soon after storms. On the other hand, skiing conditions are generally better after storms, when the slab becomes cohesive, which is necessary for slab avalanche release. In addition, slab avalanches may release on weak layers within the snowpack during times of no recent precipitation. Because of their importance, this study will focus on slab avalanches.

Figure 1.2: Skier-triggered slab avalanche (photo: B. Jamieson)

Both types of avalanches may involve dry, moist or wet snow whereas most accidents involve dry slab avalanches.

Figure 1.3: Avalanche path (photo: ASARC)

An avalanche path can be divided into the start zone, the track and the run out zone (Figure 1.3). However, not all avalanches leave the start zone of an avalanche path.

Further distinctions can be made by using various parameters including the shape of the start zone, the morphology of the track, the snow movement and the gliding layer. For a detailed explanation see McClung and Schaerer (1993: 61-89).

The release of an avalanche requires an initial failure induced by a *trigger* in order to overcome the strength in the snowpack. These triggers can be natural or artificial. Natural triggers include weather events such as snowfall, rain, temperature changes and wind, ice falls, cornice falls and earthquakes. Artificial releases include triggers by skiers (snowboarders, hikers, climbers), helicopters, over-snow vehicles, snowmobiles and explosives (CAA, 2002). Artificial triggers can be accidental or controlled, remote or nearby, or sympathetic with another artificially released avalanche. Figure 1.2 shows an accidental skier-triggered dry slab avalanche where the entry tracks can be seen in the top

right corner. The exact trigger point is unknown but was likely close to the entry tracks near the crown of the slab. The skier was carried down and partially buried. In the photograph, the skier at the crown serves as a scale for the crown thickness, which is estimated to be approximately 70 cm.

1.4 Slab avalanche failure

1.4.1 Introduction

This thesis focuses on dry slab avalanches and therefore only the failure of this type of avalanche will be discussed. McClung and Schaerer (1993: 124) defined stability as "the ratio of the resistance to failure versus the forces acting toward failure". The stability of a snowpack is largely dependent on the existence and the strength of weak layers, but also on slab properties (Schweizer and Jamieson, 2001). In recent years it has been widely accepted that a failure in shear is the primary factor for slab failure rather than a primary failure in compression or tension (Jamieson, 1995: 26-27), as was proposed in earlier studies (Haefeli, 1967). However, the shear must be fast enough to cause fracture (McClung and Schaerer, 1993: 63). The applied stresses causing failure can be natural or artificial and whether the failure occurs depends on physical properties such as density, hardness, temperature, rate of deformation and bonding to adjacent layers (McClung and Schaerer, 1993: 69).

The failure layers of slab avalanches are often distinguished into persistent and non-persistent weak layers, based on the crystal type. Persistent weak layers can be layers of surface hoar, facets or depth hoar and can remain unstable in the snowpack for weeks or months. In contrast, non-persistent weak layers such as layers of new snow crystals and decomposed and fragmented crystals typically stabilize within hours or days.

In Canada, 78% of the fatal avalanches between 1972 and 1991 occurred on persistent weak layers and only 8% on storm snow instabilities; the remaining 14% occurred on unidentified layers within the old snow (Jamieson and Geldsetzer, 1996: 17). Even though the potential for triggering storm instabilities might be higher than the percentage indicates, the reduced stability during storms is obvious and fewer people may expose themselves to avalanche terrain in stormy weather and therefore the probability of

triggering an avalanche is reduced. However, both types of snowpack instabilities will be discussed.

The high percentage of fatal avalanche accidents on persistent weak layers indicates the relevance of these layers, and the incorporation of the characteristics of weak layers into forecasting models promises to improve the forecast. However, during and soon after a storm that buries a persistent weak layer, avalanche reports do not consistently distinguish whether the avalanche ran on a persistent or a non-persistent weak layer within the storm snow. Hence models may be assessed separately for storm instabilities (persistent and non-persistent weak layers) and persistent weak layers *after* the first storm that buries them, in addition to analysing models strictly based on crystal types. The goal is to be able to run the most accurate forecasting model on a daily basis.

Natural avalanches are often associated with storm snow events and may fail on persistent and non-persistent weak layers. Although skier-triggered avalanches are the focus of this study, it is interesting to assess the predictive value of shear strength measurements for natural avalanche activity on storm snow instabilities for potential use in some forecasting operations.

1.4.2 Persistent weak layers

Persistent weak layers are the main concern for skiers with regard to avalanche accidents in the Columbia Mountains of Canada (Jamieson and Geldsetzer, 1996: 17).

Surface hoar forms on the snow surface “when the water vapour pressure in the air exceeds the equilibrium vapour pressure of ice (snow grains) at the surface” (McClung and Schaerer, 1993: 44). Figure 1.4 a shows a surface hoar crystal on the surface and Figure 1.4 b a layer of surface hoar after burial. Favourable conditions for surface hoar growth are clear, cold nights after an overcast day or in situations where there is fog close to the surface (ibid.).

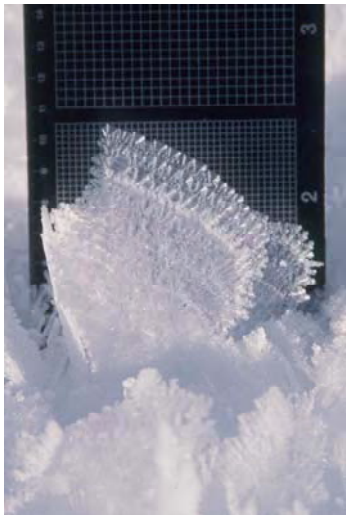

Figure 1.4 a: Surface hoar crystal before burial (photo: ASARC)

Figure 1.4 b: Surface hoar after burial (photo: ASARC)

In the Columbia Mountains of Canada, surface hoar growth during the winter months is quite common. Two to three surface hoar layers per winter can often be observed and skiers can trigger avalanches on these weak layers for weeks. Chalmers (2001) found that avalanche activity on these layers generally occurs within the first 30 days after burial.

Depth hoar, shown in Figure 1.5, is the result of strong temperature gradients, high temperatures, and large spaces between crystals (McClung and Schaerer, 1993: 49). These conditions can often be observed at the base of the snowpack.

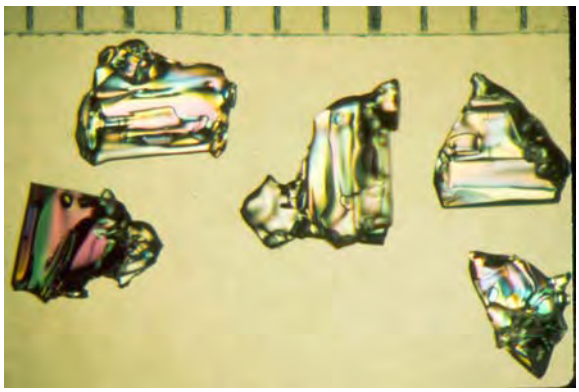

Figure 1.5: Depth hoar crystals (photo: R. Perla)

The climate of the Columbia Mountains does not favour depth hoar growth, as, for example, the climate in the Rocky Mountains does, because high temperature gradients are less often observed. Depth hoar is an advanced form of faceting (described in the next paragraph) and requires high temperature gradients over long time periods.

Facets, shown in Figure 1.6, are a focus of this study. This crystal type is formed during kinetic metamorphism (Section 1.2).

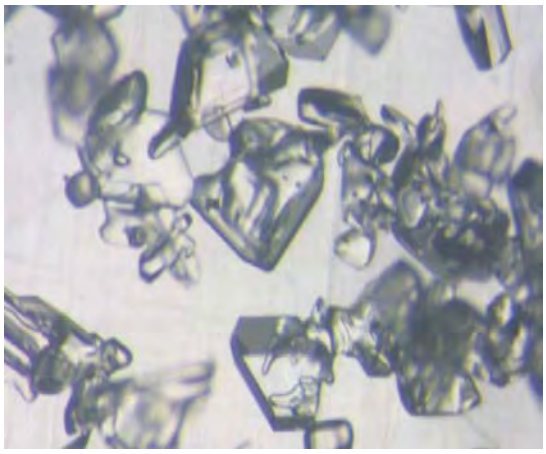

Figure 1.6: Facets (photo: ASARC)

In the International Classification for Seasonal Snow on the Ground, ICSSG (Colbeck, 1990) three forms of faceted crystals are listed. These are:

1) Solid faceted particles, which form during kinetic growth in times of high temperature gradients. The strength decreases with increasing growth rate and grain size (Colbeck at al., 1990).
2) Small faceted particles that are less than 0.5 mm in size. They usually develop near the surface through kinetic growth due to high temperature gradients. The snow is characterized by low-strength (ibid.).

3) Mixed forms are faceted crystals with recent rounding (ibid.). This occurs when the temperature gradient decreases.

This study additionally considers the subclass of *rounded grains: mixed forms* as opposed to *facets: mixed forms*. The latter occurs as a transitional form when rounded crystals develop into facets due to an increasing temperature gradient, thus decreasing the strength of the layer.

Facets may either form on the surface or within the snowpack. In this study the dataset under consideration consists of facets formed at the base of the snowpack, facets on and below crusts as well as facets formed close to the surface. The formation processes are driven by different physical conditions:

1) Facets at the base of the snowpack, hereafter referred to as November facets, typically form in November during cold time periods after major rain events. November facets can often be tracked over the whole winter and are sometimes the cause of deep slab avalanches later in the winter season. Skier-triggered avalanches can usually be observed during the first weeks after burial, but can also be of concern later when an initial failure in the upper snowpack steps down and causes this weak layer to fail, resulting in deep slab avalanches. Jamieson and Johnston (1997) investigated a November facets layer in 1996 based on weather records in two different study areas in the Columbia Mountains. They observed that the facets were weaker and slower to stabilize in areas that had less snow on the crust during the cold period in which the facets formed. The authors concluded that a thinner slab causes a higher temperature gradient and that less load does not densify the weak layer as efficiently as the higher load of thicker slabs. The November facets are usually associated with a crust, but are considered separately from facets above and below the crust as described under Point 2.
2) Facets above and below crusts are often related to avalanche activity. Colbeck (1991) assumes that facets form below crusts because the upward movement of water vapour is stopped and as a consequence the vapour pressure increases and thus provokes the growth of faceted grains. The weak layer forms above crusts when dry snow falls on the wet layer due to a high temperature gradient between

the heat source (the wet layer) and the cold air. The formation of faceted layers above crusts is explained in Colbeck and Jamieson (2001).

3) Kinetic metamorphism in the near surface snow layers is quite common but may change quickly due to temperature gradient changes and the lower densities in the upper snowpack (Birkeland, 1998, Fukuzawa and Akitaya, 1993).

Birkeland (1998) distinguishes three predominant processes that result in extreme near-surface temperature gradients and consequently in faceting. These are:

a) radiation recrystallization (faceting) due to solar radiation penetration into the snowpack and loss of heat at the surface,
b) melt-layer recrystallization due to solar radiation or rain; wet over dry, and
c) diurnal recrystallization (faceting) due to a temperature gradient as a result of warming and cooling of surface snow producing relative consistent temperatures 0.30 m below the snow.

These three processes can occur simultaneously. For further discussion see Birkeland et al. (1998).

1.4.3 Non-persistent weak layers

Skier-triggered avalanches on non-persistent weak layers are often less harmful to skiers than avalanches on persistent weak layers because they often involve low cohesive snow, thinner slabs and are generally smaller. Non-persistent weak layers such as new snow crystals (Figure 1.7 a) and decomposed and fragmented crystals (Figure 1.7 b) tend to stabilize within hours or days.

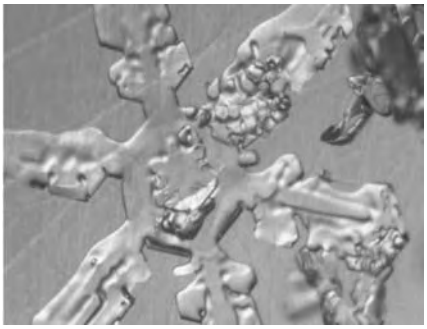

Figure 1.7 a: New snow crystal (photo: ASARC)

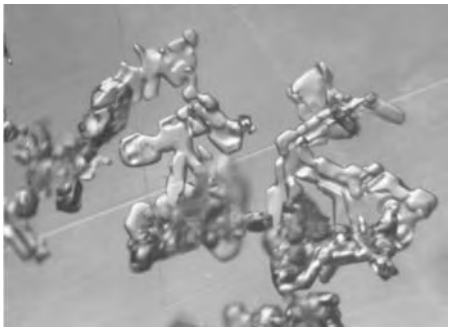

Figure 1.7 b: Decomposed and fragmented crystals (photo: ASARC)

Recent loading is the primary factor of slab avalanche releases during storm cycles (Duclos, 1998; Davis et al., 1996).

Duclos (1998) examined 22 skier-triggered avalanches in France and found that in 32% of the cases, the failure layers consisted of decomposed and fragmented crystal or new snow crystals whereas 40% occurred on depth hoar, 20% on facets and 8% on rounded grains. In Canada a similar pattern is observed with the majority of the triggered avalanches on persistent weak layers, including surface hoar.

Forecasting avalanches on non-persistent weak layers on highways and in ski areas plays a major role considering the threat of natural avalanches reaching the highway or the ski runs. Buser et al. (1985) reported that the four main variables in forecasting storm snow avalanches are new snow depth, wind (both speed and direction), new snow density and temperature, and the old snow surface (surface hoar, crusts, etc.) assuming that the old snowpack is stable.

1.5 Avalanche sizes

The *size* of an avalanche is an important indicator of its destructive potential. Table 1.1 shows the Canadian avalanche size classification (CAA, 2002), which was introduced in 1981 and is used today in most forecasting and skiing operations in Canada.

Table 1.1: Canadian avalanche size classification (CAA, 2002)			
Size & data code	**Avalanche destructive potential**	**Typical mass**	**Typical path length**
1	Relatively harmless to people.	<10 t	10 m
2	Could bury, injure, or kill a person.	10^2 t	100 m
3	Could bury and destroy a car, damage a truck, destroy a small building or break a few trees.	10^3 t	1000 m
4	Could destroy a railway car, large truck, several buildings or a forest area up to 4 hectares (~10 acres)	10^4 t	2000 m
5	Largest snow avalanche known. Could destroy a village or a forest of 40 hectares (~100 acres).	10^5 t	3000 m

Half-sizes, e.g. 1.5 are commonly recorded. Sizes of 0.5 are often recorded as signs of instability but their destructive potential is limited.

S*kier-triggered avalanches* are the focus of this study and can be summarized as avalanches that release due to the impact of a skier. Most fatal avalanche accidents in Canada are skier-triggered and the prediction of such avalanches is of great concern to skiing operations and backcountry users. Slab avalanches release on slopes with angles from about 25 to 55°, while 38° is an average value for skier triggering. The crown thickness of slab avalanches usually varies from 0.1 to 2 m (McClung and Schaerer 1993: 76) but skier triggering is most common when a weak layer is up to about 1 m deep, because the stress caused by a skier is smaller at greater depth (Föhn and Camponovo, 1997). It is difficult to forecast the likelihood of skier-triggered avalanches, because the problem is complex and requires the assessment of the avalanche danger using weather, snowpack and stability information. Even with these data available, the risk can only be reduced and not eliminated. Computer assisted forecasting models may be able to help systematically analyse the data available.

1.6 Avalanche protection

Avalanche hazard mitigation involves permanent and temporary measures to protect infrastructure, e.g. buildings and roads, or people exposed to avalanche terrain. The choice of preventative measure(s) depends on the acceptable risk for each avalanche scenario in a specific area. The acceptable risk for involuntary exposures to avalanche terrain, for example on worksites and on roads, is related to the magnitude and return period of avalanches threatening objects or people (Stethem et al., 2003). Based on a critical avalanche larger than Size 2, the acceptable return period for highways is more than 30 years. However if the expected return period is 30 years or less, permanent measures and temporary closures may be required; when avalanches of Size 2 or larger are expected with return periods of less than 10 years, active control programs with explosives are likely to be implemented (CAA, 2002). Stethem et al. (2003) state that the acceptable return period for commercial backcountry skiing is shorter than 10 years because a Size 2 avalanche already poses a danger to skiers. Again, at return periods of less than 10 years, temporary and permanent measures are applied to mitigate the hazard. For backcountry skiing operations where Size 2 avalanches are likely in most winters but permanent measures are not an option because of the amount of avalanche terrain, temporary measures are of great importance. This may involve daily stability assessments, which might affect the terrain choice or may result in non-skiing days. However, this theory is not easy to apply to voluntary exposures such as unguided ski touring in the backcountry, where the acceptable risk depends strongly on the risk perception and propensity of each individual, though mitigation measures can be offered (e.g. avalanche bulletins). An avalanche hazard with the potential to kill or injure people and cause damage to property is mitigated in Canada as summarized in Table 1.2.

Table 1.2: Avalanche hazard mitigation, after Stethem et al. (2003). Supplemented.			
		Measures	
Mitigated Area	**Potential harm**	**Permanent**	**Temporary**
Roads and railways	to traffic and passengers to economics due to delay	hazard mapping location planning defence structures	***avalanche forecasting*** detection systems temporary closures explosives control
Ski areas (in-bound)	to skiers to traffic on access roads to base facilities to ski lifts	hazard mapping location planning structural protection	***avalanche forecasting*** temporary closures ski compaction explosives control
Backcountry recreation	to skiers		***avalanche bulletins*** avalanche education weather forecasts avalanche response training
Resource industries	to buildings to roads and traffic to workers	structural defence	***avalanche forecasting*** explosives control
Energy and transmission	to transmission lines economic losses	location planning reinforced structures defence structures	
Residential and public land use	to residents and public users to property	hazard mapping building relocation defence structures support structures	evacuation plans temporary closures
Construction sites	to workers to property		***avalanche forecasting*** temporary closures evacuations plans explosives control

In most types of mitigated areas, avalanche forecasting is applied as a temporary measure for avalanche protection and could potentially involve computer models in the mitigation process. However this thesis focuses on skier-triggered avalanches and consequently is most valuable for ski areas and backcountry skiing operations.

Defence structures are generally found in the run out zones of an avalanche path and include snow sheds over roads, diversion walls for moving snow, mounds to slow the avalanche motion and reinforcement of buildings. Support structures however are generally situated in the start zones to prevent avalanche formation. The construction and maintenance of these structures can be cost intensive and are often not justifiable. The largest amount of money for avalanche protection is probably best spent where the involuntary risks and economic losses are highest. Permanent measures are more relevant for natural avalanches that may threaten highways, villages or ski runs on a specific avalanche path than for skier-triggered avalanches where the terrain choice is greatest.

Temporary measures such as bulletins, warnings, explosives control, closures and evacuations are used in most backcountry operations and in ski areas where skier triggering is of concern. Temporary measures cost less and are more flexible, while an effective avalanche warning system plays a major role (Russi et al., 2000).

1.7 Purpose of the investigation (Objectives)

Stethem et al. (2003) pointed out that better regional and local forecasting is still needed in Canada. The pressure on commercial backcountry skiing operations and forecasting services to improve their stability assessment seems to increase, especially after a catastrophic winter such as in 2002/2003. Avalanche forecasting tools such as computer assisted forecasting may become welcome tools in the stability evaluation in commercial skiing operations as well as in public warning services.

The aim of this project is to improve the forecast for skier-triggered dry slab avalanches in the Columbia Mountains of Canada where the underlying hypothesis is that the incorporation of stability indices and snowpack properties will improve the forecast significantly.

The objectives are as follows:

- **Objective 1** formulation of empirical models to calculate the shear strength of persistent weak layers on days with snowpack observations without manual tests of the shear strength of these layers
- **Objective 2** formulation of empirical models to calculate the shear strength of persistent weak layers and snowpack variables on days without snowpack observations using snowpack information from most recent profiles in representative study plots
- **Objective 3** calculation of daily stability indices using the results of the empirical models mentioned above and rutschblock tests
- **Objective 4** assess the importance of daily stability indices and snowpack properties in forecasting skier-triggered avalanches on persistent weak layers
- **Objective 5** assess the importance of daily stability indices and snowpack properties in forecasting skier-triggered avalanches on storm snow instabilities
- **Objective 6** assess the importance of daily stability indices in forecasting natural avalanches
- **Objective 7** incorporation of the daily stability indices and snowpack properties into a Nearest Neighbour forecasting model
- **Objective 8** comparing the performance of Nearest Neighbour models with and without stability indices and snowpack properties.

Figure 1.8 gives an example of how a Nearest Neighbour model is used to forecast skier-triggered avalanche activity on persistent weak layers on a daily basis.

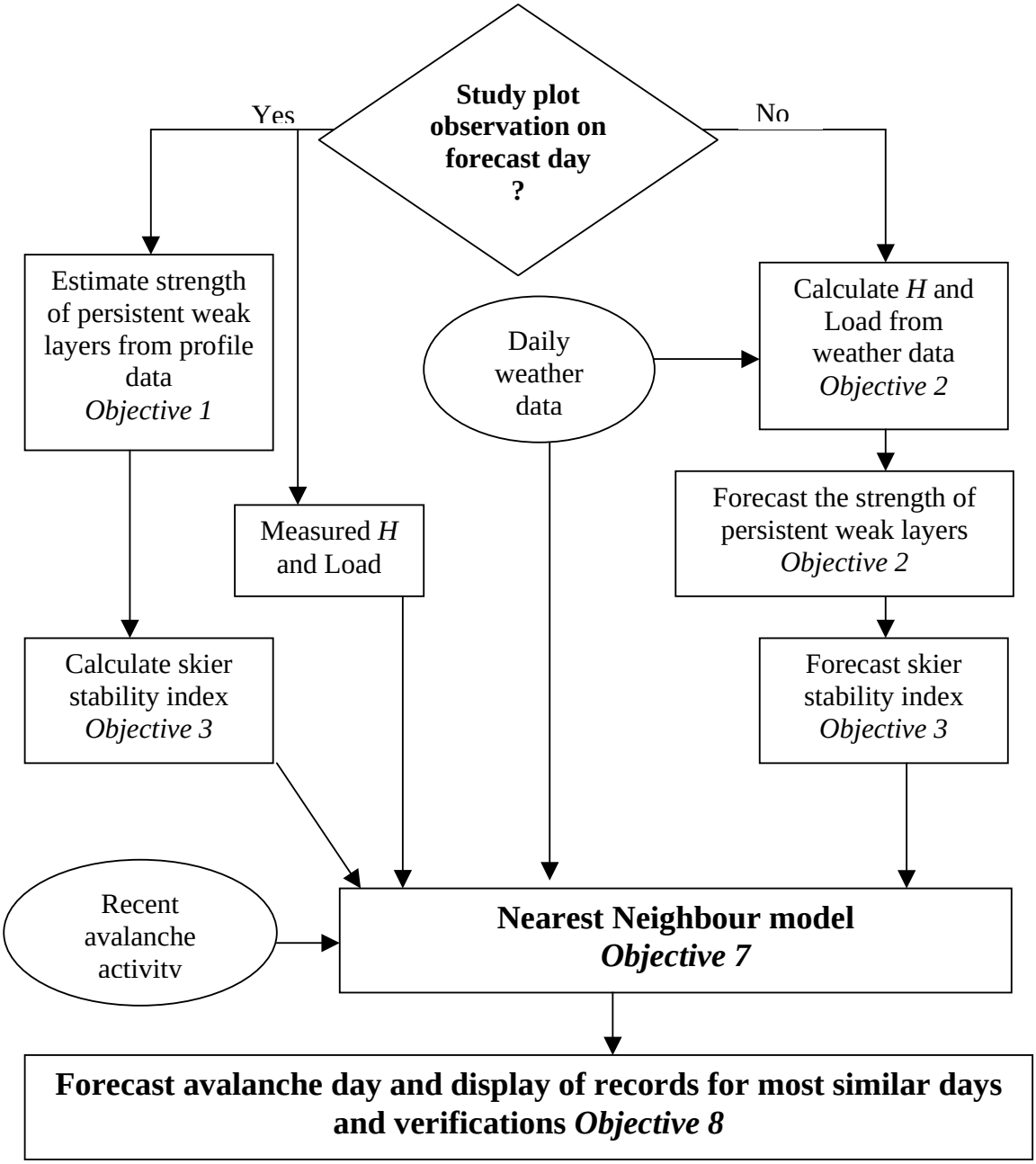

Figure 1.8: Daily use of a Nearest Neighbour model to forecast skier-triggered avalanche activity on persistent weak layers.

Because this thesis is part of an ongoing research project at the University of Calgary, parts of it involve work of previous students or an update of their analyses. Table 1.3 lists this previous work and the contributions of this thesis to the overall goal to improve the forecasting of skier-triggered avalanches in the Columbia Mountains of Canada.

Table 1.3: Previous work and contributions

Objective	Selected previous work	Contributions in this thesis
1, 2	Chalmers (2001): Forecasting shear strength and skier-triggered avalanches for buried surface hoar layers.	Shear strength adjustments to normal load; additional years of data; refined model to forecast the shear strength of surface hoar layers.
1, 2	Johnson (2000): Observations of faceted crystals in alpine snowpacks.	Shear strength adjustments to normal load; additional years of data; empirical model to forecast the shear strength of faceted layers.
3	Jamieson and Johnston (1998): Refinements to the stability index for skier-triggered dry-slab avalanches.	Shear strength adjustments to normal load for persistent weak layers.
3, 4	Jamieson (1995): Avalanche prediction for persistent snow slabs.	Stability indices calculated using the normal load adjustment. Explore the predictive value of stability indices with different statistical methods.

2. Literature review

2.1 Introduction

The aim of this thesis is to improve the forecasting by computer models for skier-triggered dry slab avalanches in order to provide a valuable tool for forecasting and skiing operations. As described in the introduction, identifying weak layers is important to assess the stability of the snowpack and to determine the potential for avalanche releases. Consequently, this literature review will review the shear strength of weak layers (Section 2.2) and its prediction (Section 2.3), including stability tests and stability indices (Section 2.4) before reviewing conventional forecasting procedures (Section 2.5). Further existing forecasting models will be introduced (Section 2.6).

2.2 Shear strength

The stability of a slab overlying a weak layer depends on the shear strength of the weak layer, which can be measured with a shear frame. Although the strength of snow can be determined by the number, size, shape and orientation of intergranular bonds (Yosida, 1963; Ballard and Feldt, 1965; Keeler, 1969; Gubler, 1982; Edens et al., 1991; Colbeck, 1997), strength changes are hard to quantify and consistent measurements are rare, thus this information is not readily available for use in computer assisted forecasting models. A good review of the bond growth theory is given in Colbeck (1997, 1998). It has been found that the bond properties largely depend on load and temperature (Keeler, 1969, Edens et al., 1991; Colbeck, 1997). Keeler (1969) reported that increasing bond area corresponds with increasing shear strength.

Colbeck (1997) defined sintering as “the process by which bonds form and the study of their size, shape, and number density”. He observed “rapidly growing grains in dry snow lack bonding, whereas strong bonds form when the grains grow slowly”. During times of rapid grain growth, rounded grains and smaller grains tend to vanish, whereas facets grow and the sintering process is limited. Keeler (1969) observed that overburden pressure is most important below 180 cm and density depends on stress rather than on time or temperature.

The shear strength is roughly proportional to density to the power 2-2.5 for dry snow, whereby facets, mixed forms and depth hoar are generally weaker at the same density than precipitation particles, decomposed and fragmented crystals and rounds. Johnson (2000) made an attempt to estimate the shear strength of a facet layers from Kojima's (1967) densification model, but he was unsuccessful. Kojima (1967) reported that depth hoar and young, new snow seem to have a lower compactive strain rate. He concludes that the strength of new snow is not only related to load.

Kojima (1967) assumed "that the strain rate of densification of a snow layer is proportional to the load exerted by the weight of overlying snow". In earlier studies (e.g. 1954) he and Feldt and Ballard (1966) used accumulation rates constant over time whereas in his later research Kojima (1967) introduced a "time-integrated load" into his equation, which is possibly more accurate for strength changes over time. Chalmers (2001) also developed a load lagged model for the shear strength of buried surface hoar layers.

Keeler and Weeks (1967) pointed out that few studies were done to monitor mechanical properties, including strength measurements over time and that most of the earlier studies had been done on higher density polar snow (350 kg/m^3) rather than on less dense snow. "Measurements of the strength properties of mountain snow have been much more sporadic, particularly those coupled with detailed pit observations" (Keeler and Weeks, 1967). Since 1993 the University of Calgary research group (ASARC) have consistently observed strength changes over time, including snowpack observations, shear frame and stability tests. In Switzerland some studies have been done on low density snow using tests from soil mechanics (Bader et al., 1939; Bucher, 1948; de Quervain, 1950) and in Japan, Yoshida (1955) used more refined techniques. The focus of the earlier studies was the correlation of strength and temperature (Roch, 1965) and strength and density (Keeler and Weeks, 1967). Keeler (1969) found a good correlation between strength and density in low density snow. Because it became obvious that the grain type was of major importance, most studies distinguished between fine grained, wet and depth hoar crystals (Keeler and Weeks, 1967) whereas the importance of weak layers on avalanche activity had not yet been pointed out. Compared to layers of precipitation particles, decomposing and fragmented particles and rounded grains, layers of faceted grains exhibited increased

variability of tensile and shear strength (Sommerfeld, 1973; Jamieson and Johnston, 1990, 2001).

Keeler (1969) related his work strongly to the work of Bader et al., (1939) and Yosida (1955-58) to determine the spatial variation of the physical properties of snow. He stated that density correlates with the mechanical strength properties of dry snow, but not with wet snow and depth hoar.

Johnson (2000) related the shear strength of weak layers of faceted crystals in the Columbia Mountains (mostly Intermountain snow climate) and Rocky Mountains (Continental snow climate) to easily measured snowpack properties. He found positive correlations with load, slab density, slab thickness, hardness of the facet layer, snowpack thickness and age of the facet layer, while load was the primary snowpack variable that affected the shear strength of layers of faceted crystals. In the first four to six weeks after the layers formed, the increase in shear strength was roughly proportional to the loading rate in each of the two snow climates. Jamieson and Johnston (1997) monitored a particular layer of November facets throughout the winter in 1996 in the Columbia Mountains, including shear frame measurements. The authors calculated two stability indices, SF and RB calc and correlated the indices with avalanche activity (natural and skier-triggered). They found that these indices might not be as valuable for deeper weak layers, possibly because they did not use a normal load adjustment as will be discussed in Chapter 5.

The prediction of the shear strength of recently deposited weak layers that release storm snow avalanches over time is difficult, because such weak layers may form and stabilize within hours and shear strength measurements are generally not done on an hourly basis.

2.3 Forecasting the shear strength of weak layers

Running a forecasting model on a daily basis requires daily values of the predictor variables, which are not readily available. Therefore forecasting models have made little use of shear strength measurements of weak layers. Unfortunately shear strength measurements, e.g. with a shear frame, are time consuming and are not done on a regular basis in most forecasting operations. This is a disadvantage, because forecasters could use

this information in the decision making process. But little research has been done to forecast the shear strength of weak layers.

For data from the Columbia Mountains of western Canada, Chalmers (2001) and Chalmers and Jamieson (2002) regressed current shear strength and the rate of shear strength change for surface hoar layers on measurements from manual snow profiles. When these two regression models were combined, estimates of shear strength and stability within eight days of a snow profile were promising. In these studies the normal load effect on the shear strength measurements with the shear frame was not considered but might improve the shear strength prediction for deeper burials.

2.4 Stability measurements and stability indices

2.4.1 Introduction

The assessment of the stability of the snowpack is crucial in determining whether avalanches are likely or not. However, evaluation of stability is a complex task and requires interpretation of information available from snowprofiles, snowpack tests, and avalanche observations. Only a few tests give direct information on the stability of a snowpack, including shear frame (Sommerfeld, 1984), rutschblock (Föhn, 1987b), compression (CAA, 2002), stuffblock (Birkeland and Johnson, 1996), and shovel shear tests (CAA, 2002). The importance of the site selection and the experience of a person interpreting the results effectively need to be emphasized.

"Most tests identify potential weak layers or interfaces and give an index of snowpack stability" (Schweizer et al., 2003). All but the shear frame tests involve properties of the weak layer and the slab, and, except for the shear frame and shovel tests, include dynamic surface loading, as is the case for skiers (Schweizer and Jamieson, 2001). The shear frame test is independent of the properties of the overlying slab, because the slab is taken off to a few centimetres above the weak layer. However, due to the removal of the slab, the normal load of the slab during testing is neglected, although it can be adjusted when calculating the shear strength. Föhn and Camponovo (1997) showed that results from the shear frame test and the skier stability index are correlated.

In this study the shear frame test and the rutschblock test are important and are used to estimate a skier stability index of a weak snowpack layer. The techniques are described in Section 4.2. The shear strength of a weak layer is used to calculate stability indices as a quantitative indicator for the forecaster. These indices have been developed over the last 40 years to better predict avalanche activity, including refinements due to various triggers. For example, Föhn (1987a) introduced a skier stability index, which accounts for the shear stress induced by a skier. In addition, indices have been refined to extrapolate the results from a slope to a regional scale. Sommerfeld (1976) and Perla (1977) showed that a stability index can be related to slab avalanches, however they did not then distinguish between natural and artificial triggers. However, these indices are calculated based on point measurements, and spatial variability of snowpack properties may compromise extrapolation from the study plot to the regional scale (Hägeli and McClung, 2001; Landry et al., 2002). It has been shown that on a slope scale the variability of test results in avalanche start zones can be high (Campbell, 2004), but it has also been shown that test results correlate to avalanche activity on a regional scale when the test site is chosen carefully (Jamieson, 1995; Föhn, 1987b; Jamieson, 1999).

2.4.2 Shear strength measurements

Shear frame tests are most frequently used to measure the shear strength of weak layers in a snowpack. As will be described in detail in Section 4.2.2, a shear frame is placed parallel to the identified weak layer and a few mm above it. The overlying slab is removed and a force gauge attached to the shear frame is pulled quickly until the weak layer fractures. The maximum force at failure is recorded (Sommerfeld, 1984). The shear strength is calculated as the maximum force divided by the shear frame area. Other techniques used for testing soils, such as the shear vane, have proved impractical for shear strength measurements in the snowpack, because placement a few millimetres above the weak layer is difficult (Jamieson, 1995: 27). The shear frame test assumes that the failure is first in shear, and loading to failure within one second ensures brittle fracture.

The shear frame test, as with all other shear strength tests, has its flaws, some of which can be corrected. Perla (1977) and Sommerfeld (1984) reported a shear strength

decrease with increasing frame area. Sommerfeld and King (1979) adjusted the shear strength for size effects calculating the Daniels strength, the shear strength of an arbitrarily large failure area.

Another influence during testing is that the slab is removed, “thereby reducing the normal load during testing” (Jamieson and Johnston, 1998). Roch (1966) suggested an adjustment for different grain types, which was used in 1998 by Jamieson and Johnston for non-persistent weak layers, although they assumed that the effect was negligible for persistent weak layers in their dataset. However, more detailed analysis has shown that an effect seems to exist for all weak layers. Consequently, in this thesis an approach is made to determine the normal load adjustment for persistent weak layers, which promises to improve the shear strength estimations, especially for deeper burials.

A detailed description of the factors affecting the shear frame test results, such as the frame size, the frame type, the normal load and different operators can be reviewed in Jamieson (1995: 65-106).

In addition to shear strength measurements with the shear frame, the results of the rutschblock test will be used to estimate the stability index Sk38 of a persistent weak layer (Section 7.3.4).

2.4.3 Stability indices

In general, stability indices are calculated as the ratio of strength to stress. Even though stability indices have been calculated since the 1960s, the interpretation has had to be refined over the years. As mentioned above, some limitations in the shear frame test had to be assessed and different triggers analysed. In addition, shear frame measurements are point measurements and most valuable for avalanche forecasting purposes when measured in a uniform study plot and extrapolated to the surrounding terrain. The results were verified over the years by comparing the test results with observed avalanche activity. However, in most of the studies the stability indices were compared to slopes on which avalanches had already released (e.g. Conway and Abrahamson, 1984), instead of also comparing the indices to slopes on which no avalanche released, as was done later by Föhn (1987a) and Jamieson (1995: 147-155).

One of the earliest stability indices is the stability factor (SF) (Schleiss and Schleiss, 1970), which is also known as the stability ratio (CAA, 1995). SF is the shear strength of a weak layer divided by the weight per unit area above the weak layer (load). It has been used in stability assessments since 1962 by the avalanche control section at Rogers Pass to predict natural avalanche activity on storm snow instabilities at a highway corridor in Glacier National Park. An SF index of < 1 indicates low stability, SF between 1 and 1.5 transitional stability and SF > 1.5 good stability. However SF is not considered a real stability index because it is not a ratio of shear strength to shear stress.

Another common stability index for natural slab releases is S, which was introduced in 1966 by Roch and is based on the ratio between shear strength and shear stress due to the slab. Roch adjusted the shear strength to the normal load effect for various grain types. Föhn (1987a) combined normal load adjustments and shear frame size adjustments (Sommerfeld and King, 1979; Roch, 1966) to calculate S as

$$S = \frac{\Sigma_\infty + \sigma_{zz}\phi}{\sigma_{xz}} \qquad \textit{Equation 2.1}$$

where Σ_∞, the Daniels strength, is calculated as

$$\Sigma_\infty = 0.56\Sigma_{100} \qquad \textit{Equation 2.2}$$

$$\Sigma_\infty = 0.65\Sigma_{250} \qquad \textit{Equation 2.3}$$

where Σ_{100} is the shear strength measured with a 0.01 m^2 shear frame,

Σ_{250} the shear strength measured with a 0.025 m^2 shear frame,

σ_{zz} the normal load in the form of

$$\sigma_{zz} = \rho g h \cos^2 \psi \qquad \textit{Equation 2.4}$$

where ρ is the density of the slab (kg/m^3), g is the acceleration due to gravity (9.81 m/s^2), h is the slab thickness (m) and ψ is the slope angle,

ϕ is the normal load adjustment and σ_{xz} is the shear stress in the weak layer due to the weight of the overlying slab in the form of

$$\sigma_{xz} = \rho g h \cos^2 \psi \qquad \textit{Equation 2.5}$$

Roch (1966) analysed the effect of normal load on weak layers of precipitation particles, rounded grains and depth hoar. However, in this study only his equation for precipitation particles will be used:

$$\phi(\Sigma_\infty, \sigma_{zz}) = 0.08\Sigma_\infty + 0.056 + 0.022\sigma_{zz} \qquad \textit{Equation 2.6}$$

Jamieson and Johnston (1998) analysed the normal load effects for surface hoar and layers of faceted crystals, but none of Roch's adjustments fitted their Canadian data. However, their dataset was limited and because they could not find a dominant normal load effect, they assumed ϕ = 0 for persistent weak layers. In the following studies by various researchers at the University of Calgary, the shear strength of persistent weak layers was not adjusted for the normal load. In Chapter 5 an attempt is made to recalculate the shear strength adjustment for persistent weak layers.

2.4.4 Refined indices for skier-triggered avalanches

Föhn introduced the first artificially triggered stability index in 1987a to improve the information on slope instability in regard to observed avalanches. He introduced various equations for skiers, climbers and snowcat triggers. The general form originates from Equation 2.1 and the term $\Delta\sigma_{xz}$ was introduced in the denominator as the shear stress induced by a potential trigger:

$$\Delta\sigma_{xz} = \frac{2R\cos\alpha_{max}\sin\alpha_{\max}{}^2\sin(\alpha_{\max} + \psi)}{\pi h \cos\psi} \qquad \textit{Equation 2.7}$$

where R is the line load due to a skier and α_{max} is the angle from the snow surface to the peak shear stress in the weak layer.

For a skier, Föhn (1987a) assumed a weight of 85 kg and a static line load over 1.7 m. The equation for skier-triggered avalanches (S') is:

$$S' = \frac{\Sigma_\infty + \sigma_{zz}\phi(\Sigma_\infty, \sigma_{zz})}{\sigma_{xz} + \Delta\sigma_{xz}} \qquad \textit{Equation 2.8}$$

By distinguishing between the human triggered and natural index, Föhn (1987a) could improve the forecast of 110 avalanches from 55% to 75% using shear frame measurements on the particular avalanche slopes.

However, his equation assumes that the skier's weight is applied at the surface of the snow and does not take the ski penetration into account. Therefore Jamieson (1995: 141) adjusted the term $\pi h \cos\psi$ in the denominator of Equation 2.7 to $\pi(h - P_k)\cos\psi$ where P_k is the ski penetration during skiing and can be estimated in the Columbia Mountains as:

$$P_k = 0.55 - 0.0016\rho_{30} \qquad \textit{Equation 2.9}$$

where ρ_{30} is the slab density at 30 cm

This consequently increases the skier-induced stress on the weak layer. The modified skier stability index has the form:

$$S_k = \frac{\Sigma_\infty + \sigma_{zz}\phi(\Sigma_\infty, \sigma_{zz})}{\sigma_{xz} + \Delta\sigma'_{xz}} \qquad \textit{Equation 2.10}$$

where σ'_{xz} has been adjusted for ski penetration.

Jamieson and Johnston (1998) improved the forecast over S' in their study by using their new adjustments.

Föhn (1987b) related the stability index S' with the rutschblock results. He found that the first three rutschblock scores related mostly with stability levels of S' less than one. In the same study the author calculated S and found that "the mean values of S are clearly

larger than those of S', i.e. the probability of natural slab releases is two to three times smaller than the probability for slabs triggered by skiers at a given rutschblock degree."

Even though most of the shear strength measurements required for the calculation of stability indices are from shear frame measurements, the results of other stability tests such as the rutschblock and the compression test can be used to estimate stability indices. Jamieson (1995: 178-179) found the relation of Sk and the rutschblock score as Sk = 0.31(RB-1) and because Sk is a function of shear strength, the shear strength of a weak layer can be estimated from rutschblock test results.

2.4.5 Refined regional stability indices

Most of the indices mentioned above are used to assess the stability on a slope scale and are based on point measurements. The results thus depend on the exact slope inclination in the specific start zones, whereas extrapolated indices can be applied to start zones with various slope inclinations (Jamieson, 1995: 36) assuming that the shear strength was measured in a representative study site. By calculating S for a 35° slope, a typical slope angle for dry slab avalanches in the Columbia Mountains, Jamieson and Johnston (1993) introduced S35 as an extrapolated stability index. Their index successfully predicted avalanche activity on 75% to 87% of 70 days. They also calculated a skier-stability index Sk35, but Jamieson (1995: 125) found 38° as a better typical angle of slab avalanche start zones. He found that an Sk38 value of less than 1.5 indicates instability on persistent weak layers; however, more avalanches released at values of Sk38 less than 1.

Sn38 is the extrapolated stability index for natural avalanches based on Equation 2.1 and calculated for a 38° slope. Sk38 is based on Equation 2.10 and calculated for a 38° slope. Both indices will be used as input variables for forecasting models in this thesis.

In the Columbia Mountains, Sk38 correlated with skier-triggered avalanches within 100 km of the study plot in which the snow profile was observed (Chalmers and Jamieson, 2003). Zeidler and Jamieson (2004) successfully incorporated a skier stability index into a daily forecasting model to predict skier-triggered avalanches. They used Chalmers' (2001) work to extrapolate the shear strength of the surface hoar layer over time based on study plot observations and on Zeidler and Jamieson (2002) who used a similar approach to

forecast the shear strength of faceted layers. However, in both studies the shear strength was not adjusted for the normal load effect. Consequently in Chapter 6 the extrapolations of shear strength for persistent weak layers will be recalculated including data from three additional years.

2.5 Avalanche forecasting

2.5.1 Introduction

In general, avalanche forecasting is the prediction of current and future snow stability (McClung and Schaerer, 1993: 164). If the destructive potential is considered, the term avalanche hazard forecasting is applied (CAA, 2002). Often the assessment of the avalanche hazard is relevant in forecasting operations. For example, avalanches potentially reaching a highway are of concern and need management – a factor critical to the forecasting program. Smaller sized avalanches that are not considered a risk to motorists may be used as indicators of the stability of the snowpack.

Well established avalanche control programs, e.g. on highways and in ski areas, primarily use conventional avalanche forecasting methods, as described in LaChapelle (1980) to assess the stability of the snowpack. The decision making process involves the evaluation of given parameters based on experience, intuition, and the local knowledge of the avalanche forecaster. However, in more recent years, models have been developed to help the forecaster predict snow stability and thus to provide support but do not provide a definite forecast. In the end, the forecaster is responsible for assessing the risk and deciding whether precautions such as road closures or explosives control are appropriate or not.

2.5.2 Conventional avalanche forecasting

LaChapelle (1980) describes the fundamental processes of conventional avalanche forecasting as a mix of deterministic treatment for snow and weather parameters and inductive logic to reach actual forecast decisions based on the forecaster’s experience. Weather, snow structure, avalanche occurrence (past and present), test skiing, artificial release, and local climate history all contribute to the forecast (LaChapelle, 1980). The variables used for operational forecasting depend on the availability and reliability of these

variables, their relevance in a specific forecasting area, the forecasting programs' purposes, time of the year and the reasoning and experience of the forecaster. As said by LaChapelle (1970) there is more than one way to forecast an avalanche. Correlated predictor variables are often available, which allow individual successful forecasters to prefer different (but correlated) predictors. Point stability tests such as compression tests and rutschblock tests are done at various times and locations within the forecast area; however, these test results are important in the decision-making process.

McClung and Schaerer (1993: 125) distinguished between three classes of forecasting factors:

> **Class I:** Stability factors, including current avalanche activity, stability tests (e.g. rutschblock, compression and shovel tests) as well as other signs of instability, such as cracking of the snowcover.
>
> **Class II:** Snowpack factors, such as the structure of the snowpack, including the existence of weak layers, crystal forms and sizes, densities and snow temperatures.
>
> **Class III:** Meteorological factors, such as the temperature, winds, precipitation and humidity.

The uncertainty of these factors increases with the class number – this means that Class I factors are more direct indicators of snow instability than are those of Class II and Class III. However, in more recent years the entropy in stability tests due to spatial variability became apparent so that it was proposed to reclassify stability tests into Class II or Class I b. In conventional avalanche forecasting, forecasters use the information of Class I and Class II factors effectively by applying their knowledge over time, even though snow profiles are generally observed only once a week and stability tests are done intermittently and often at varying locations.

Different forecasting operations deal with different scale issues when predicting stability or avalanche hazards. In general, the more local and detailed the forecast, the more important are Class I and Class II factors, whereas forecasting natural avalanches on a regional scale may rely mainly on good weather information. LaChapelle (1970) associates predicting *direct action* avalanches, which occur during or soon after storms, with using

meteorological methods, whereas structural methods examine the shear strength and layer patterns to identify the hazards; the emphasis shifts with the climate but also possibly with the trigger. As discussed previously, weak layers play a major role in determining the avalanche hazard for skier-triggered avalanches, which often occur between storms.

Even though storm snow avalanches are thought to be best forecast using meteorological predictor variables (Ferguson at al., 1990; LaChapelle, 1970), in Chapter 10, Class I and Class II factors will be analyzed to assess their predictive value for natural avalanches. Also, although Hägeli and McClung (2001) postulate that snowpack characteristics are of more value locally and only representative of a smaller area, in this thesis snowpack and stability indices will be used to forecast skier triggering on a regional scale.

2.5.3 Computer assisted avalanche forecasting

In this thesis, computer assisted forecasting is defined as a forecast in which a computer is used in the decision making process at any time. Here the simplest form may be the graphical visualization of weather and snowpack information as time lines, snow cones, snow roses, snow profiles, etc., which are described in Atkins (1992). Besides data visualization, computer assisted forecasting may involve data management, data analysis and data exchange. Computer programs can help to handle the amount and complexity of data relevant in decision-making processes by systematically exploring the relationships between variables and their analysis of the variation over time and in space. The quantity and the quality of the response and predictor variables determine the output possibilities of the models.

While in conventional avalanche forecasting Class I and Class II factors are used over time based on the experience of the forecaster, computer models require daily values of input variables and consequently the data available on an irregular basis have to be extrapolated over time.

Although computer models have been developed since the 1970s (Judson and Erickson, 1973; Bois et al., 1974; Salway, 1979) the acceptance and the demand has only increased over the past several years. The basic ideas behind these models have not

changed, except that due to better computing capacities the interfaces are becoming more user-friendly, better visualization techniques are available and it is easier to configure these models to a specific forecasting area. Up to now forecasters seem to most appreciate the data visualization such as the graphical representation of forecasting factors and the mapping of avalanche activity over the forecast area. However, data analysis tools have great potential to be valuable for daily avalanche forecasting although forecasters have to learn how to use the information models provided effectively. Existing forecasting models will be discussed in more detail in the next section.

2.6 Avalanche forecasting models

2.6.1 Introduction

As mentioned previously, forecasting models can be more than visualization and data management tools. Indeed, these tools can help analyze complex data patterns. Analytical methods include deterministic, statistical, and/or expert systems, which are well known and researched (Schweizer and Föhn, 1996). Some of these models are used in forecasting services; others are presently at the research stage.

2.6.2 Deterministic models

Deterministic models try to simulate the snowcover processes with physical formulas. In hydrological sciences, numerical snowpack models are used for calculating the run-off in spring, but for avalanche forecasting purposes it is not enough to know the water equivalent of the snowpack. In addition, it is crucial to model the layered characteristics of the snowpack, including weak layers along with the grain types, grain sizes, temperatures, and densities. CROCUS (Brun et al., 1989, 1992), developed in France, and SNOWPACK (Lehning et al., 1999), developed in Switzerland, are the snowpack models currently most relevant to avalanche forecasting. CROCUS models a typical snowpack at different locations at varying elevations and aspects whereas Lehning et al. (1999) criticize this approach because it does not take the spatial variability on a small scale and the wind distribution into account. However, the authors accept the applicability on a regional scale (Lehning et al., 1999). As a result SNOWPACK tries to model the local snowpack at many

distinct sites in order to capture spatial variability (Lehning et al., 1999). Although downscaling procedures and drift models were incorporated in SNOWPACK, the success in regard to spatial variability has not yet been achieved. Both snowpack models still have difficulties in modelling weak layers, which are crucial especially for skier-triggered avalanches (Fierz, 1998). The simulated snowpack properties can be used for forecasting purposes or they can be further included into statistical models or expert systems. In Switzerland, for example, the forecasting service uses calculated new snow settlements from remote weather stations in their decision-making process. In France, the snowpack structure, as modeled by CROCUS, is further used in the expert system MEPRA, which estimates the avalanche hazard.

2.6.3 Statistical (data based) models

Statistical models use a historical database to find the most significant weather and snowpack variables related to avalanche activity in a specific area and give an indication of the stability. The most common methods include discriminant analysis, nearest neighbour analysis and classification tree analysis.

Judson and Erickson (1973) describe *discriminant* analysis as a multivariate statistical technique for assigning data into two or more groups based on prior knowledge. Logical discriminations in avalanche forecasting are avalanche day (yes/no), avalanches (none/dry/wet) or the size of an avalanche. This analysis can also be used to assess the importance of predictor variables.

The basic premise of a *Nearest Neighbour Model* is that similar snow and weather conditions lead to similar avalanche situations (Schweizer and Föhn, 1996). Days with historical values of meteorological and other variables that are close to values for the forecast day are identified as nearest neighbours; avalanche activity is forecast that is similar to that reported on the nearest neighbours (days). Even though a probability can be calculated, the list of the avalanche activity on days with similar conditions is more meaningful to the forecaster (Buser, 1989). Commercially available Nearest Neighbour Models, for example NXD and Cornice, only predict avalanche day (yes/no). McClung and Tweedy (1994) use parametric discriminant analysis using Bayesian statistics and cluster

techniques in discriminant space to analyze avalanche occurrences by the method of Nearest Neighbours. The authors use the Mahalanobis distance, instead of the Euclidean distance, to account for the correlations of predictor variables in their non-parametric discriminant analysis (Nearest Neighbours calculations). The two distance calculations will be discussed in Section 4.3.2. McClung and Schaerer (1993: 165) criticize the nearest neighbour approach, because "there is no formal way to test whether a variable is really contributing to a proper result or whether it is redundant".

Tree analysis is used to find classifiers and to explore the interaction between predictor variables (Breiman et al., 1984). The classes can be defined in accordance with the forecasting purpose, e.g. the maximum size of an avalanche (Davis and Elder, 1995; Jones and Jamieson, 2001; Zeidler and Jamieson, 2004).

2.6.4 Expert systems

Expert systems try to simulate the reasoning of an avalanche forecaster. Through symbolic computing it is possible to incorporate information available to a forecaster, for example days on which control work was already done, which cannot be put into the computer model in mathematical terms. These models are flexible in the input but do not rigorously calculate a specific result. Schweizer and Föhn (1996) developed MODUL and DAVOS in Switzerland; they use an external judgment processor and have "trained" it to recognize specific avalanche situations. The hazard is forecast as an output. Because some of the input is manual and the programs are interactive, these models can be time consuming to run, which might be a disadvantage in operational forecasting. MEPRA, developed in France, assesses the snow stability and deduces a risk of natural avalanches and of skier-triggered avalanches (Durand et al., 1999) by combining deterministic (CROCUS) and statistical and heuristic methods (Giraud, 1992). McClung's (1995) snow profile assistant is an expert system, which assists persons with only a basic level of experience to analyze snow profiles. Attempts have been made to integrate neural networks into avalanche forecasting systems, which have the advantage of using incomplete and inconsistent data (Stephens et al., 1995). Schweizer et al., (1994) introduced ALUDES as an avalanche forecasting system that integrates neural networks and rule-based systems and

determines rules to assess the avalanche danger. AVALOG is another interesting decision support system based on expert knowledge and was introduced by Bolognesi (1992). This model is a real time system and the output is a map of the avalanche hazard. The verification of AVALOG was promising, though the author stated that it could be improved by including statistical methods.

2.7 Relevance to objectives

Some models try to predict the hazard levels while others try to determine the likelihood of an avalanche occurrence. The models are becoming more and more flexible and easier to adapt to specific forecasting purposes. As mentioned earlier, on highways it might be more important to forecast the size of an avalanche to determine the potential of reaching the highway, whereas skiing operations are more interested in the potential of skier triggering, where even a relatively small slab avalanche might be harmful. Many researchers (Buser et al., 1985) point out that a meaningful variable for "avalanche activity" is not available. They find that "the often-used 'avalanche-day-probability' yields only the likelihood for a day with at least one avalanche, which is difficult to translate for warning purposes" because it does not specify the number of avalanches or *where* avalanches are likely to occur.

The incorporation of stability indices and slab properties into a Nearest Neighbour Model promises to improve the avalanche forecast, especially for skier triggering. Some approaches have been made to incorporate the shear strength of the snowpack into forecasting models as for example in France, where the snowpack simulations of CROCUS are used in MEPRA to deduce additional mechanical characteristics such as shear strength and rammsonde resistance and to add this information to the snowprofiles (Durand et al., 1999; Merindol et al., 2002). In the Columbia Mountains of western Canada, Chalmers and Jamieson (2003) developed an empirical model based on shear frame measurements to forecast the shear strength of surface hoar layers up to eight days ahead using snow profile data on days with observations. An update of their model will be developed in Section 6.4. In addition, a similar empirical model will be introduced in Chapter 6.3 to forecast the shear strength of faceted layers.

2.8 Spatial analysis in avalanche forecasting

In the past years spatial statistics have become more sophisticated and some avalanche research studies have applied spatial analytical techniques (Stoffel et al., 1998; Laternser and Schneebeli, 2002; McCollister, 2004). However most studies are based on the use of Geographic Information Systems (GIS) to visualize avalanche activity. McCollister (2004) went further and combined a meteorological nearest neighbours technique with a GIS. He used three weather variables to investigate the avalanche activity for a ski area and for subregions for similar aspects and elevation and for individual avalanche paths with success (McCollister, 2004: 3).

Attempts to apply geostatistical kriging to avalanche data in Switzerland were unsuccessful. Laternser and Schneebeli (2002) tried to find a spatial pattern based on climate and avalanche activity of various locations; however they rejected kriging because the variograms showed only a random pattern, meaning that closer locations were not more similar than the once further apart. They suggested that the data quality of the different stations might have had a great influence on the results. In a different study on the regional scale in the Columbia Mountain of Canada (Zeidler, term paper) the variograms showed also only a random pattern. On a slope scale neither Campbell (2004) nor Stewart (2002) could find good variograms for most arrays of stability tests indicating that the stability is not more similar in nearby locations. This reflects the spatial variability of stability.

Considering the quality of the data, the extent of the study area and the aim to provide a day-by-day avalanche forecast spatial analysis is not the technique of choice in this thesis. Laterneser and Schneebeli (2002) pointed out that high data quality is required to analyse the avalanche activity on a larger scale. However in this study the focus is on skier-triggered avalanche activity, which occurs less often than natural avalanches and whether avalanches are triggered or not depends mainly on the amount of skiing done on one day. This means that when the snowpack is very unstable and no skiing was done or due to good terrain choices of the guides no skier triggering occurred on one day but triggering was likely, no avalanches were reported. This information is crucial for daily avalanche forecasting models. Another difficulty is the large study area where individual avalanche paths are not observed continuously and not every elevation and aspect might be

skied on each day. Schweizer and Kronholm (2004) observed that weak layers, which are important to forecast skier triggering, varied spatially within a basin and adjacent terrain. A forecast based on aspect and elevation on a regional scale might not account for this and therefore the regional forecasting for skier-triggered avalanches in this thesis is based on the worst-case scenario, meaning that the likelihood of avalanches *anywhere* within the forecast area is assessed and not as much the likely locations of avalanches.

Anyways the nearest neighbour model does imply some spatial information in the list of the avalanche occurrences and it would be useful to link this information to a GIS, which would help the forecaster to determine where (e.g. elevation, aspect) avalanches are likely. The forecaster may be able to interpret the outcome of the nearest neighbour model in a spatial sense with the list of the avalanches on the ten nearest neighbour days. Further, mapping may become useful in the future but that is beyond the scope of this thesis.

3. Study areas and datasets

The Columbia Mountains of western Canada are the main study area for this thesis. They are located west of the Rocky Mountains and east of the Coast Mountains. Three main snow climate zones are generally distinguished: maritime, transitional and continental (LaChapelle, 1966; Armstrong and Armstrong, 1987, McClung and Schaerer, 1993; Mock and Kay, 1992). More recently Hägeli and McClung (2003) analyzed the avalanche characteristics in the Columbia Mountains in relation to the local snow climate and they defined the Columbia Mountains as having a transitional snow climate with a strong maritime influence. LaChapelle (1966) mentioned this type of snow climate as being coastal transitional and situated between the coastal and intermountain snow climates. This distinction was not included in the three main climate zones, but has great importance in western Canada (Hägeli and McClung, 2003).

LaChapelle (1966) described the snow climates with the associated avalanche characteristics as follows:

- *Maritime snow climate*: heavy snowfall, mild temperatures; mainly new snow instabilities, which may stabilize quickly due to mild temperatures; rain possible throughout the winter, which may cause widespread avalanche cycles.
- *Continental snow climate*: typical in the Rocky Mountains; low snowfall and cold temperatures; shallow snow covers; favourable climatic conditions to form persistent weak layers, which stabilize slowly.
- *Transitional (intermountain) snow climate*: combination of maritime and continental influences; deep snowpacks with some persistent weak layers.

The maritime influence in the Columbia Mountains may vary significantly from year to year and thus result in the existence of persistent weak layers. Indeed, Hägeli and McClung (2003) found that the percentage of natural avalanche activity on persistent weak layers varies between 0% and 40%, depending on the maritime influence.

Data from two study areas in the Columbia Mountains of western Canada were mainly used in this project. The first study area is situated near Blue River (Figure 3.1, Area 1) and the second in Glacier National Park (Figure 3.1, Area 2).

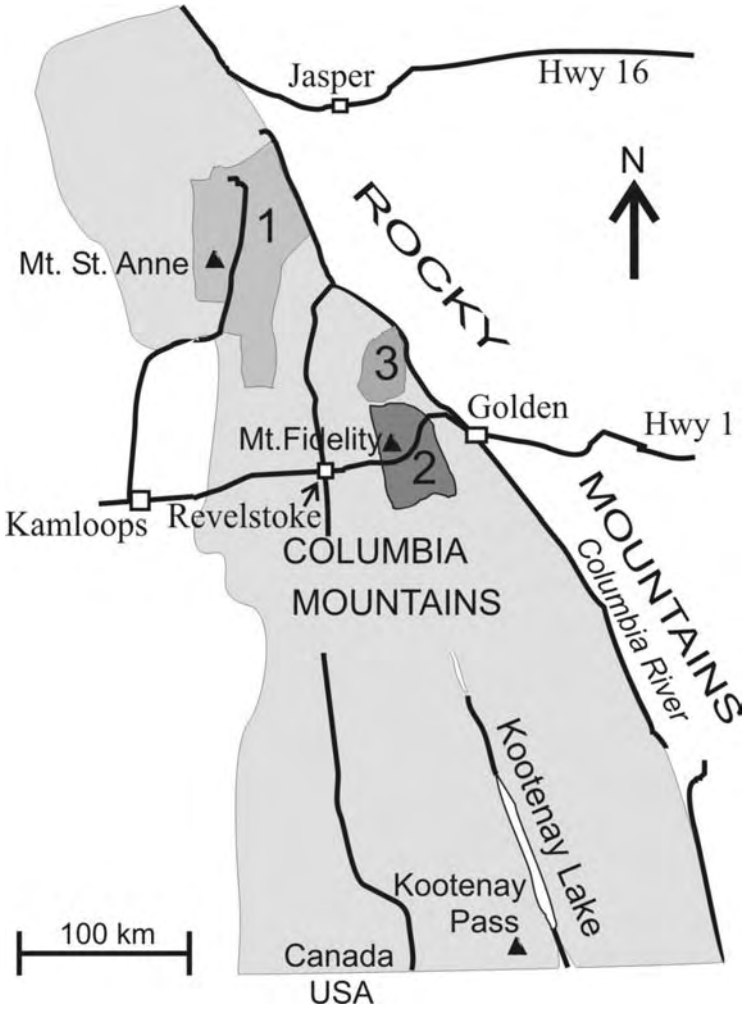

Figure 3.1: Map showing the location of study sites at Mt. St. Anne and Mt. Fidelity and areas in which avalanche activity is recorded and used in this study. 1: Mike Wiegele Helicopter Skiing (MW), 2: Glacier National Park (GP), 3: CMH Adamants (AD).

Meteorological variables were available from remote weather stations at Mt. St. Anne at an altitude of 1900 m near Blue River and at Mt. Fidelity at an altitude of 1905 m as well as Round Hill at an altitude of 2010 m on Mt. Fidelity in Glacier National Park. Snowpack data and the strength of persistent weak layers were recorded once or twice per

week in a level study plot on Mt. St. Anne near Blue River and on a study slope (25-37°) on Mt. Fidelity by field workers from the Applied Snow and Avalanche Research Group (ASARC) at the University of Calgary. Additionally Mike Wiegele Helicopter Skiing (Figure 3.1, area 1) provided avalanche occurrence data. The avalanche control section in Glacier National Park GNP (Figure 3.1, Area 2) and CMH Adamants (Area 3) provided avalanche occurrence data, meteorological and snowpack data. A few datasets from the Rocky Mountains and the Bobbie Burns (BB) study plot, located to the southeast of Glacier National Park, were used in Chapter 6. Table 3.1 gives an overview of the data from the different areas used for each objective and the years in which the stability measurements and snowpack and avalanche observations were taken. The author was involved in the fieldwork during the winter 2001/2002 – 2003/2004 at ASARC GNP.

Table 3.1: Data used by location and objective.

Objectives	ASARC RP	ASARC MW	MW	AD	BB	GNP	Rocky Mountains
1, 2	1999 - 2004	1993 - 2004	-	-	1993 - 1998	-	1999 - 2000
3, 4	1995 - 2004	1995 - 2004	1995 - 2004	1995 - 2004	-	-	-
5	-	-	1991 - 2004	1995 - 2004	-	1995 - 2004	-
6	-	-	-	-	-	1995 - 2004	-
7, 8	-	-	1995 - 2004	1995 - 2004	-	1995 - 2004	-

This project was only possible due to consistent measurements over a long time range, and thus the author was not involved in all the years of fieldwork.

4. Methods

4.1 Introduction

In Chapters 1 and 2, the layered snowpack characteristics and the importance of the shear strength of weak layers and snowpack stability were discussed. The interpretation of these snowpack characteristics along with the results of shear strength measurements and stability tests help to assess the stability of the snowpack. In Section 4.2.1 a typical snowprofile, in Section 4.2.2 a shear frame test and in 4.2.3 a rutschblock test will be described as the most important field methods used in this study. In addition to these field observations, classification tree analysis will be outlined in Section 4.3.1, because this statistical method might not generally be known, though it has proved to be a valuable technique in some studies of avalanche forecasting (Davis et al., 1996; Jones and Jamieson, 2001; Davis and Elder, 1995; Rosenthal and Elder, 2002). Furthermore the basics of the Nearest Neighbour analysis (Section 4.3.2) will be introduced along with the main features of Cornice (Section 4.3.3), a Nearest Neighbour avalanche forecasting model developed in Scotland. In the last section of this chapter (4.3.4) forecast verification techniques will be described.

4.2 Field methods

4.2.1 Snowprofile observations

Figure 4.1: Snowprofile (photo: ASARC)

A standard snowprofile (Figure 4.1) in a study plot (level study site) or on a study slope (sloping study site) was typically observed every two weeks as a full depth snow profile and additionally once or twice a week as a test profile to the depth of a particularly interesting persistent weak layer. Both the full and the test profile generally included:

- the height of the snowpack
- the thickness of each layer
- the grain form for each layer
- the grain size for each layer
- the hand hardness for each layer
- the density for each layer
- the liquid water content for each layer, and
- the temperature of the snow typically measured at 10 cm intervals from the surface.

The observations were done according to the Canadian observation guidelines for snowprofiles (Canadian Avalanche Association, 2002) with three exceptions:

1) the thickness of weak layers was measured to the nearest millimetre at three or more places on the pit wall and averaged
2) the load (overburden) was the average weight of cylindrical samples from the snow surface to the weak layer divided by the cross-sectional area of the cylinder, and
3) the temperature of the snow was additionally measured 5 cm above, 5 cm below and within a previously identified weak layer.

Figure 4.2 shows a simplified snowprofile typical for the Columbia Mountains.

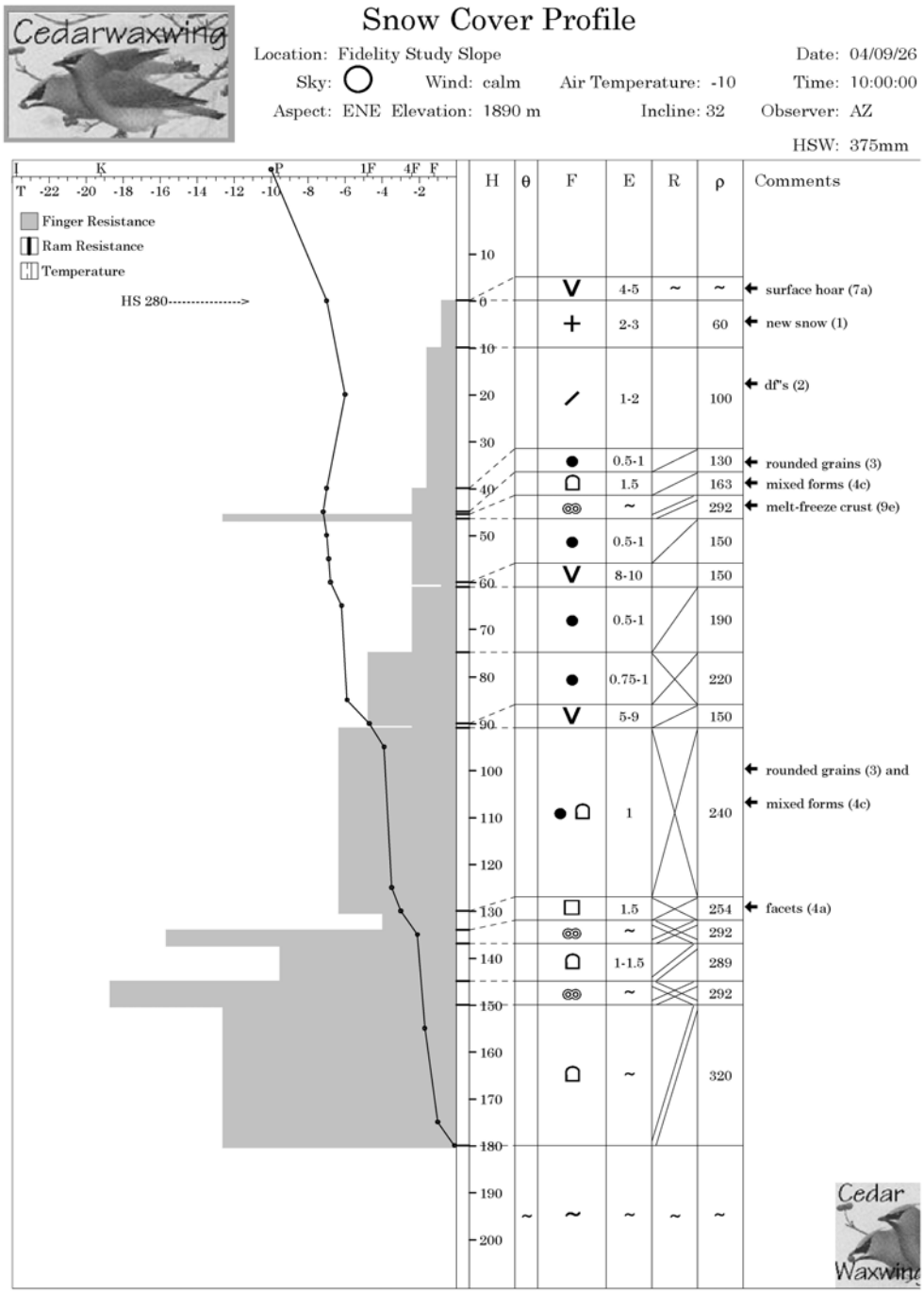

Figure 4.2: Manual snowprofile plotted with the snow profile program Cedar Waxwing. Crystal forms (F) are explained in the comments. E is the crystal size, R the resistance and ρ the density (CAA, 2002).

Most of the layers in the snowpack can be identified visually, whereas other layers can only be determined by differences in hardness or mechanical tests such as the rutschblock test. Most major weather events during the winter can be reconstructed from the layers in the snowpack, e.g. rain events form crusts and cold weather periods form layers of surface hoar or facets. As a general rule it is often said that the more homogenous the snowpack the more stable it may be and that a thin snowpack with lots of different layers is an indicator of unstable snowpack conditions. The identification of snowpack layers and their properties is important, but the difference in hardness and grain size across the layer boundaries are important for the stability assessment (Schweizer et al., 2003).

4.2.2 Shear frame test

The shear strength of weak layers was measured with a shear frame as described in Sommerfeld (1984) and Jamieson and Johnston (2001) generally once or twice a week, depending on the age of the weak layer (Figure 4.3). Older layers tend to gain strength more slowly than younger layers and consequently less frequent measurements were necessary for older layers. Testing of a weak layer stopped after 30 days for surface hoar layers, whereas testing on facets continued until the layer was not identifiable any more or the fractures were not planar.

The general procedure involves the identification of the weak layer(s) currently tested in the snowprofile or by previously performing stability tests such as compression or rutschblock tests. Persistent weak layers were named according to their date of burial.

Testing with a shear frame requires that the overlying snow be removed so that the frame is filled when pushed into the slab and the bottom is about 2-5 mm above the weak layer to be tested. This promotes planar shear fractures. Because snow adjacent to the layer might bond to the frame, a blade was used to cut around the frame and loosen these bonds. A force gauge attached to a cord on the shear frame was pulled smoothly and quickly (< 1 sec) to cause a brittle failure in the weak layer. The maximum force (kPa) was recorded along with the description of the fracture surface (Table 4.1).

Table 4.1: Descriptor for fracture surface	
C	Smooth, planar fracture surface
SBD	Small back divot (< 5 mm in depth)
MBD	Divot under rear compartment, 5-10 mm deep
BBD	Big back divot (> 10 mm in depth)
BC	Divot extends beyond rear compartment
W	Wavy, more than one wave corresponding to frame sections (5-10 mm)
SW	Slightly wavy (wave height < 5 mm)
LC	Lateral chunk – fractured deeper at left or right side
IRR	Irregularities 5-10 mm deep
SIR	Irregularities < 5 mm deep
SH	Small hump (< 5 mm in height)
H	Hump (> 5 mm). Reject test, if frame pops up noticeably
STP	Fracture steps between two planes < 5 mm apart

It is important to note the irregularities, because the forces required to fracture the weak layer might be influenced by non-planar fractures. In the statistical analysis some outliers were associated with the non-planar fractures. An average of 12 shear frames per weak layer was pulled, which reduced the standard error (standard deviation of the mean).

The shear frame size was either 0.01 or 0.025 m^2. The larger frame was preferred, but at times where the larger frame was hard to place because the slab was either very soft or very hard the smaller frame was an advantage.

Figure 4.3: Measurement of the shear strength of a weak snowpack layers with a shear frame

The shear strength is the average maximum force of the twelve measurements divided by the frame area.

Sources of error such as the shear frame size and the normal load adjustment were discussed in Section 2.4. Additionally the stress concentration due to the shear frame may influence the result, because the shear stress in the weak layer is concentrated below the cross-members that subdivide the frame, and at the front and back of the frame. When the frame is placed in the weak layer the stress concentration is increased and lower strength values measured (Jamieson and Johnston, 2001). Different operators can also be a source of variable results because of their pulling techniques. Spatial variability in the study plot or on the study slope can be reflected in fluctuating strength values, for instance pockets of well preserved surface hoar crystals may result in lower strength values than expected for most of the slope. If the study plot measurements are used to assess the stability on a regional scale they have to develop weak layers that are representative of start zones in the forecast area. These effects are summarized in Jamieson (1995: 27.32) and Jamieson and Johnston (2001).

4.2.3 Rutschblock test

Rutschblock tests were done according to the Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches (CAA, 2002). Rutschblock tests (Figure 4.4) provide an indication of the stability of the snowpack on a slope relevant to skier triggering, though the test site has to be chosen carefully so that an effective interpretation is possible. The selected site should be safe, undisturbed and representative of the avalanche terrain under consideration. The dimensions of a rutschblock test are 2 m across and 1.5 m upslope when the side walls are shovelled and 2.1 m across the front of the block and 1.9 m across the back when a saw or a rope is used to cut the side walls. This tapering prevents friction on the site walls, which might lead to false scores. In Figure 4.4 the sites were shovelled. The back wall can be cut with a rope, a saw or the tail of a ski.

Rutschblock tests were done during regular study plot observations as well as on different sites throughout the study area.

Figure 4.4: Rutschblock test

The rutschblock scores according to the Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches (CAA, 2002) are:

RB1: The block slides during digging or cutting.

RB2: The skier approaches the block from above and gently steps down onto the upper part of the block (within 35 cm of the upper wall).

RB3: Without lifting the heels, the skier drops from straight leg to bent knee position, pushing downwards and compacting surface layers.

RB4: The skier jumps up and lands in the same compacted spot.

RB5: The skier jumps again onto the same compacted spot.

RB6: For hard or deep slabs, remove skis and jump on the same spot. For soft slabs or thin slabs where jumping without skis might penetrate through the slab, keep the skis on, step down another 35 cm, almost to mid-block and push once then jump three times.

RB7: None of the loading steps produced a smooth slope-parallel failure.

Campbell (2004) showed that rutschblock scores may vary on one slope significantly and that it is possible that an RB1 score is right beside an RB7. This implies

that a single result has to be interpreted carefully. The slope angle is important and Föhn (1987b) suggests an inclination of at least 30°. However, Jamieson and Johnston (1993) found that shallower slopes to about 25° are good indications for the slope stability. They also reported an approximate increase of one rutschblock score for each 10° decrease in slope angle.

In addition to the score, the fracture character is recorded as proposed in van Herwijnen and Jamieson (2004), though incorporation of fracture character into computer assisted forecasting models has not yet been attempted. Also the release type, i.e. whether the whole block, part of the block or only an edge released, is not assessed as a predictor variable in computer models, but could possibly improve the forecast.

Rutschblock results have to be interpreted carefully based on experience and terrain knowledge. One slope can be highly variable and a decision based on only one test may be costly; consequently the importance of the site selection should not be underestimated.

4.3 Statistical methods

4.3.1 Classification tree analysis

Classification trees are used to determine the importance of predictor variables associated with the response variable when meteorological and snowpack variables are used in combination, and to understand the interactions.

Classification trees consist of nodes that are each split into two subsets using a critical value of one variable, as seen in Figure 4.5, where the first node, including all data points, is split into two subsets using a critical value of Sk38 ≤ 1.33.

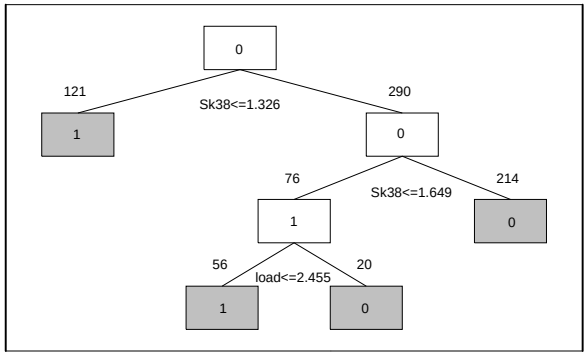

Figure 4.5: Example of a classification tree

The resulting child nodes are split until they become terminal nodes (grey boxes), either when only cases belonging to the predicted class are present in the node or a node contains a minimum number of cases as defined by a *stopping rule*. In the analysis throughout this thesis stopping occurs when there are five or fewer cases in a node unless otherwise stated. Applying stopping rules in classification tree analysis prevents the tree from growing to a perfect fit for the dataset as low-order splits would have little predictive value for other datasets based on similar conditions (Jones and Jamieson, 2001).

Discriminant-based univariate splits are used to determine the best splitting variable. For each predictor variable a significance level (*p*-value) is calculated and the variable with the lowest *p*-value is used to split the cases in a node into two subsets.

The *right-sized tree* should be as simple as possible, but should be sufficiently complex to account for patterns in the data likely to occur in similar datasets. *Minimal cost-complexity cross-validation pruning* was chosen, which is based on the misclassification error. Here the originally grown tree is pruned upwards by deleting branches of the tree until the child node is reached. A 3-fold cross-validation was used where the dataset is split into three samples of which two in turn are used to calculate auxiliary classification trees and the third sample is used as a test sample to predict the accuracy of the computed classification trees. For each of the subtrees, the cross-validation costs and errors are calculated following Breiman et al. (1984: 66). The subtree with the lowest classification cost is defined as the right sized tree.

The *misclassification cost* is generally the proportion of misclassified cases. The advantage of calculating the misclassification cost instead of looking at the misclassified cases is that it is possible to consider that a more accurate prediction for one class is sometimes desired. Setting *priors* on each class influences the cost calculation, and with this, the development of the tree. In this thesis the focus is on correctly classifying instability more often than stability and the priors were set accordingly. Setting a higher prior on unstable days tends to decrease the misclassification rate for these cases (Breiman et al., 1984: 112).

As mentioned, cross-validations use the number of misclassified cases to estimate

the probability of misclassifying a case (Breiman et al., 1984: 73). *Global cross-validation* was used to test the tree performance. The classification tree analysis is replicated a specific number of times, in our case three times, each time withholding a fraction (1/3) of the dataset, which is used as a test sample to cross-validate the tree. In the end, each of the three cases is tested once in a classification tree. The misclassification cost and error are calculated, giving an indication of the tree performance. Additionally an independent test dataset was used in most cases in this thesis to test the model's accuracy.

Further, importance ranking of the predictor variables was performed. Here all predictor variables were ranked in accordance to their potential effect on the classification (Breiman et al., 1984: 41, 147). This is important because the final tree structure alone may often not be indicative of the potential importance of the predictor variables.

4.3.2 Nearest neighbour analysis

Nearest Neighbour models are the most popular statistical models for forecasting avalanche activity. The most similar days (Nearest Neighbours) based on values of meteorological and other variables to the day to be forecast are determined. The avalanche activity, weather and snowpack observations on these days are listed and interpreted by the forecaster.

It is general practice to scale the variables by standardizing them before the distances are calculated so that each variable has the same influence on the distance calculation (Manley, 1994: 60). In addition, most Nearest Neighbour avalanche forecasting models allow weighting of predictor variables in accordance with their importance.

Most avalanche forecasting models use the Euclidean distance metric to determine the distance between *p* variables as

$$d_{ij} = \sqrt{\sum_{k=1}^{p} (x_{ik} - x_{jk})^2} \qquad \textit{Equation 4.1}$$

where *i* represents a forecast day, *j* represents a day in the historical database, and *x* is a predictor variable.

However the Mahalanobis distance accounts for correlations between variables and is calculated as

$$D_{ij}^2 = \sum_{r=1}^{p}\sum_{s=1}^{p}(\mu_{ri} - \mu_{rj})v^{rs}(\mu_{si} - \mu_{sj}) \qquad \textit{Equation 4.2}$$

where v^{rs} is the element in the r^{th} row and s^{th} column of the inverse of the covariance matrix for the *p* variables (Manley, 1994: 63). When the population covariance matrix is the same for all populations, this measure can be calculated.

The Euclidean distance is often criticized because it does not take into account that variables may be highly correlated and essentially measure the same thing (Floyer, 2003: 67, McClung and Tweedy, 1994). Floyer (2003: 67) clearly favours the Mahalanobis distance because it is sounder statistically, though he compared forecasts using both distance measures and found the results to be similar, with a fit of about 71% for avalanche/non-avalanche days (Floyer, 2003: 79). He further argued that though it is easier to modify weights using the Euclidean distance the Mahalanobis distance would be more objective, because weighting is difficult and does not allow the forecasters to set weights towards more recent years which are fresher in their memories. However by precluding weighting, the forecaster's knowledge of instability patterns, and interaction of terrain and local weather are not considered. Perhaps for this reason, the two most popular nearest neighbour models NXD2000 (Gassner et al., 2000) and Cornice (Purves et al., 2003) use the Euclidean distance. Cornice, as described in the next section, uses an automatic weighting scheme by default, which ensures an objective weighting over the whole dataset, though Cornice is based on the Euclidian distance. However it is possible in Cornice to set the weights manually, allowing the forecaster more flexibility. Most other Nearest Neighbour models set the weights according to the forecasters' experience. Purves (personal communication, 2004) justifies the use of the Euclidean distance based on the fact that the measure is easy to explain to the forecaster and therefore more easily accepted whereas although the Mahalonobis distance is statistically more correct, it is not as comparable to conventional forecasting procedures.

A comparison between the two distances is not the aim of this project, which is rather to assess the role of snowpack variables in an available model and to provide a useful tool. Correlations between the snowpack variables are discussed for multivariate methods: classification trees and nearest neighbour model.

Figure 4.6 shows a simple example of the concept using the Euclidian distance and two predictor variables: precipitation and wind speed. All observations in the dataset are represented either as dots (days without avalanches) or crosses (days with avalanches). T stands for today (the forecast day) and the ten most similar days are circled. In the graph, avalanches were reported on eight out of ten days for the Nearest Neighbour of the T in the upper right hand corner. The probability forecast is 80% and an avalanche day would thus be forecast for the day since most Nearest Neighbour models forecast an avalanche day when three or more of the Nearest Neighbours were avalanche days (categorical forecast). Heierli et al. (2004) has already pointed out that all probability forecasts have their value but the value depends on the intended application and the underlying data.

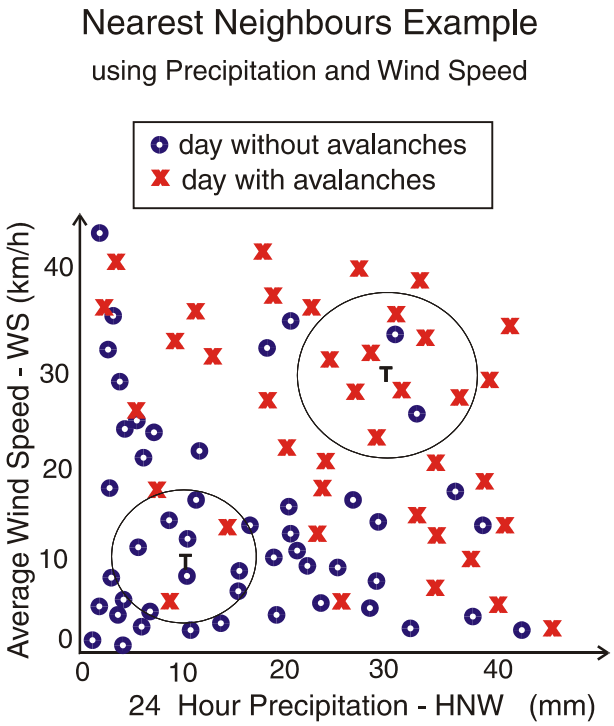

Figure 4.6: Avalanche forecast using the Nearest Neighbour approach (Graph: B. Jamieson).

This is only a simple example and in models used operationally, the number of predictor variables varies significantly and depends on the availability of the predictors and also on the data quality.

The forecaster can interpret the results based on her experience and local knowledge. Purves et al. (2003) found that the Nearest Neighbour method is well accepted by forecasters because the operation is intuitive and similar to the process of conventional forecasting.

4.3.3 Cornice

Cornice (Purves et al., 2003), a Nearest Neighbour model developed in Scotland, is used in this study as software to forecast avalanche activity. The model is operationally used in some forecasting areas in Scotland and is well accepted by the forecasters. The developers focused strongly on its practical application. The performance of Nearest Neighbour models depends strongly on the availability of significant input variables, their scaling and weighting as described earlier.

As with most of the commercially available Nearest Neighbour models, Cornice uses the Euclidean distance metric to determine the most similar days. However, instead of calculating the distance between input variables on a forecast day, Cornice considers the three previous days in order to predict patterns of similarity (Purves et al., 2003). This is an advanced method and has to be kept in mind when comparing this method to, for instance, the results from the classification tree analysis.

The raw input data has different ranges and Cornice implements an automatic scaling procedure to normalize the data ranges. Scaling of variables simply means that all variables are set to have the same range, allowing a less biased distance calculation. Cornice transforms each parameter to the same range through the use of its maximum and minimum values (Purves et al., 2003). However, scaling by using standard deviations is common practice in other Nearest Neighbour models such as NXD2000, developed in Switzerland (Gassner et al., 2000).

Weighting of each input variable determines its importance. The weighting in Cornice can be done through the use of a genetic algorithm (Purves et al., 2003) or by setting them manually using the local knowledge of the forecasters or results of prior statistical analysis. NXD2000 requires weighting variables manually and consequently prior analysis is essential for each forecasting area. The weighting is done by trial and error to maximize the model's performance. Purves et al. (2003) see the advantage of the automatic weighting function in optimizing the model's performance over the entire dataset. They found that a manual configuration may bias the model towards recent winters or serious memorable events, because it is difficult for a forecaster to set the weights in a multi-dimensional space. A detailed description of the generic algorithm can be found in Purves et al. (2003). However they warn that the automatic weights cannot be interpreted as having a physical meaning.

The number of Nearest Neighbours considered could vary from model to model, though Cornice uses ten Nearest Neighbours and the threshold for forecasting an avalanche day is fixed at three. This means that an avalanche day is forecast as when three or more of the Nearest Neighbours were avalanche days. This implies that avalanche days are more important to forecast but also reflects that most of the datasets are unbalanced as regards the ratio of avalanche days to non-avalanche days. However if the dataset becomes more balanced, the results may be influenced towards forecasting avalanche days more accurately and frequently. Floyer (2003: 70-73)) analysed balanced and unbalanced datasets in regard to the optimum number of the Nearest Neighbours considered and the optimum threshold and found there was a difference for balanced and unbalanced datasets, though the prediction accuracy was about the same. For an unbalanced dataset he found 30 Nearest Neighbours with a threshold of 7 days to be best, however his dataset was larger than most of the datasets used in this study and some of 30 Nearest Neighbours may already be quite distant from the forecast day. Nevertheless, the ratio 7:30 is similar to 3:10 as used in Cornice.

The output of Cornice consists of a list of ten Nearest Neighbours including information on the predictor variables and avalanches (Figure 4.7). In the top row the information from the forecast day is given, where the third column represents the avalanche

code. The 1022 stands for an avalanche day and 0 for a non-avalanche day. This information is entered at the end of the day for the forecast day. In this example one out of the ten Nearest Neighbours was reported as an avalanche day and the remainder as non-avalanche days. Consequently the categorical forecast in the morning predicted a non-avalanche day with a probability of 10% for an avalanche occurrence. On this specific day the forecast was wrong, as at the end of the day an avalanche was reported. Figure 4.8 shows this information graphically.

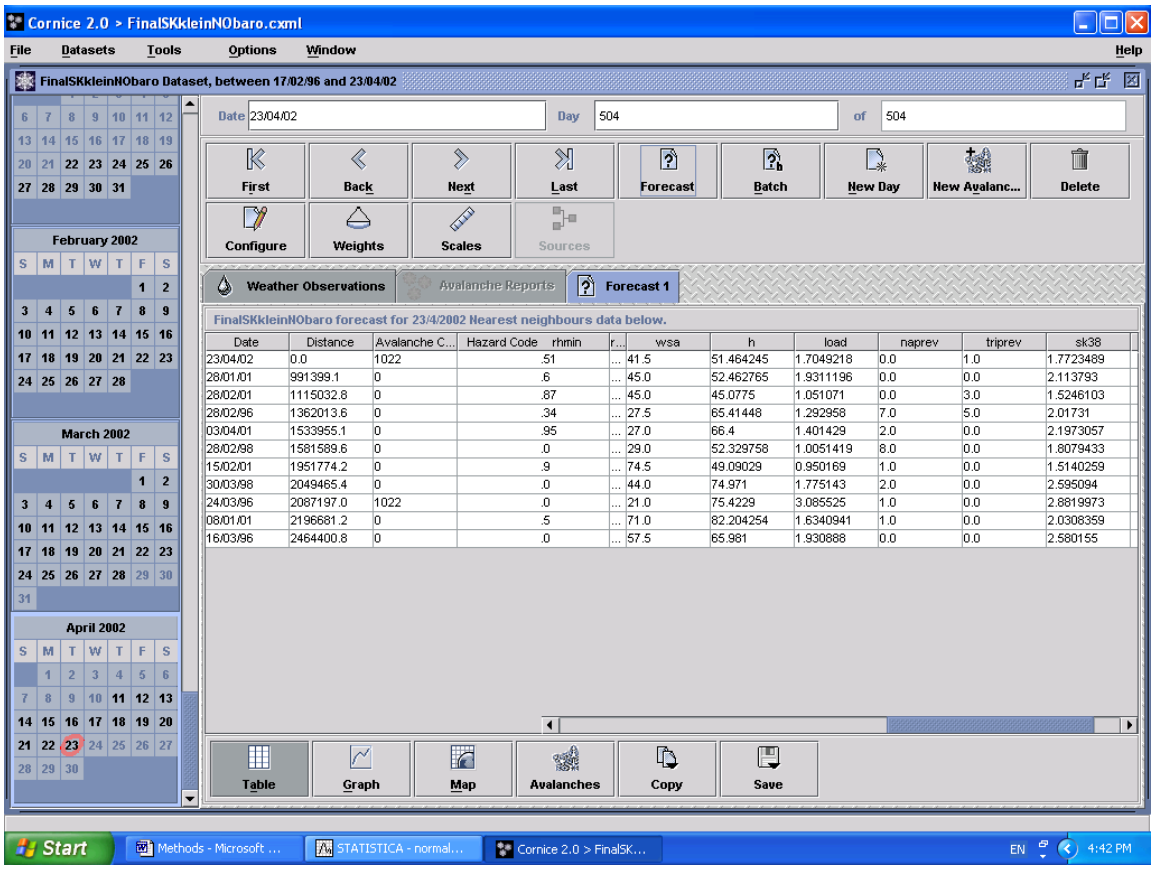

Figure 4.7: Forecast of Cornice (descriptive)

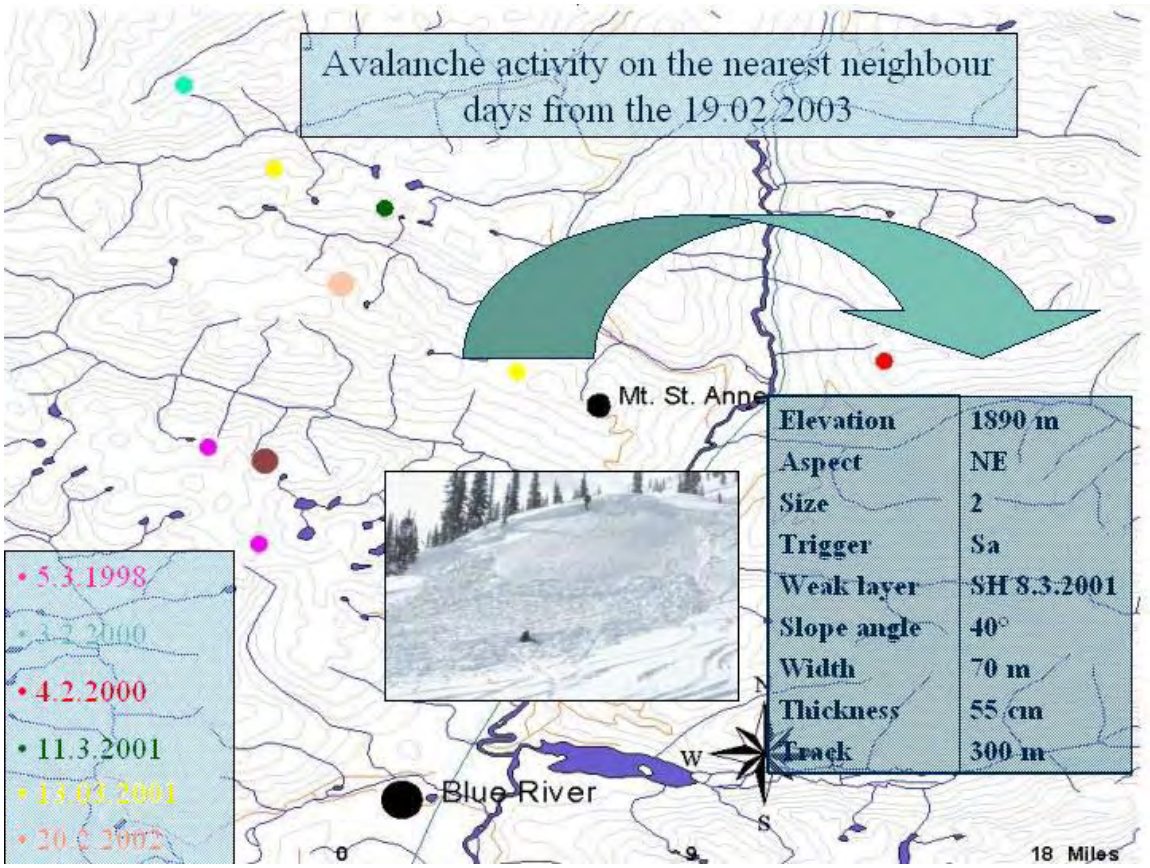

Figure 4.8: Example of the display of the avalanche activity on the Nearest Neighbour days. The dots represent avalanche activity and the information available on each occurrence.

A so-called batch forecast is implemented in Cornice to test the performance of the whole dataset. A forecast is considered correct if three or more of the most similar days have avalanches and on the forecast day there was avalanche activity or if fewer than three of the most similar days indicate avalanches and no avalanches occurred. The number of correctly forecast avalanche and non-avalanche days is displayed as well as the overall percentage (Figure 4.9).

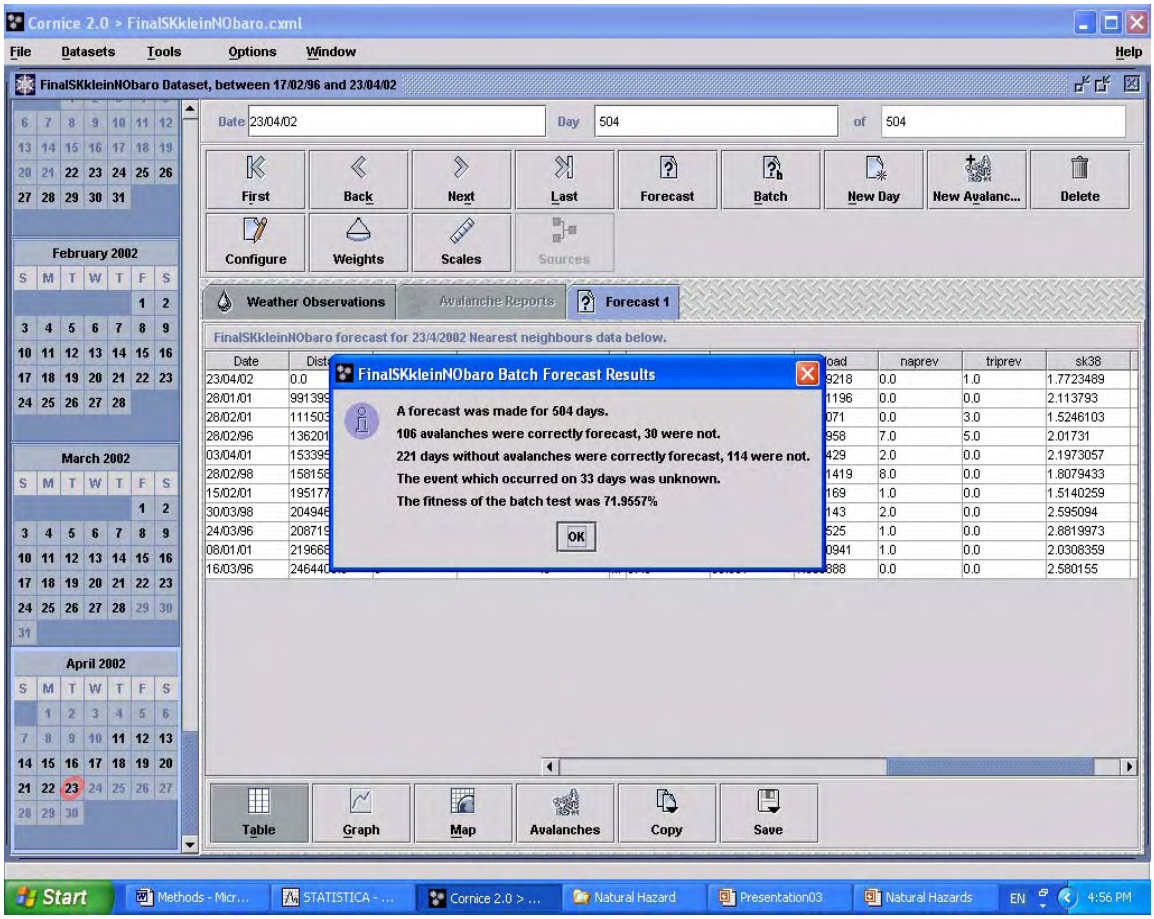

Figure 4.9: Output of a batch test of Cornice

This is one possible testing of the model, however it is not an independent test because it uses the learning sample as a test sample, but even though the performance may be overstated it is one indicator of the model's performance. In this thesis a test and a separate learning sample will be considered to enable independent testing. The performance of a Nearest Neighbour model generally improves with the size of the dataset. In addition to this verification it is general practice to calculate skill and accuracy measures from contingency tables as will be described in the next section.

4.3.4 Forecast verification

The results of forecasting models have to be verified in order to assess the quality of the forecast (Wilks, 1995: 233). In avalanche research it is common to verify categorical forecasts (e.g. avalanche day yes/no), avalanche probability forecasts and descriptive forecasts such as lists of avalanche activity in Nearest Neighbour models. Each forecast is compared to the actual event that occurred on the forecast day.

Of particular interest in this thesis, are dichotomous forecasts (avalanche days - yes/no). The verification techniques can involve accuracy measures and skill measures. Accuracy measures compare individual forecasts with the events they predict and a single number describes the overall quality of a set of forecasts (Wilks, 1995: 236). Forecast skill is the relative accuracy of a set of forecasts compared to a reference forecast, where a skill score represents a percentage improvement over the reference forecast (Wilks, 1995: 237).

The simplest way of displaying a categorical forecast is a contingency table of absolute frequencies, which represents the possible combinations of forecasts and actual events. Table 4.2 shows such a contingency table for the avalanche day (yes/no) forecast.

Table 4.2: Contingency table
A-D = the number of forecast/observed pairs

N = a + b + c + d		Observed	
		Yes	No
Forecast	Yes	a	b
	No	c	d

In terms of avalanche forecasting, a represents the number of correctly forecast avalanche days, b avalanche days forecast when no avalanche occurred (false unstable), c forecast non-avalanche days when an avalanche occurred (false stable) and d, correctly forecast non-avalanche days. The forecast of false stable days usually has higher consequences than the forecast of false unstable days. Consequently it is more important to

avoid incorrectly forecasting non-avalanche days than avalanche days. However, forecasting too many false unstable days may lead to mistrust in the avalanche forecasting model.

The most straightforward accuracy measure is the hit rate, which is calculated as

$$H = \frac{a+d}{n} \qquad \textit{Equation 4.3}$$

However the hit rate, as pointed out by Wilks (1995: 240), does not distinguish between correct yes and correct no forecasts. Consequently the calculated score is equally affected by false stable forecasts as it is by false unstable forecasts.

One variation of the hit rate is the unweighted average accuracy (UAA) calculated as

$$UAA = 0.5\left(\frac{a}{a+c} + \frac{b}{b+d}\right) \qquad \textit{Equation 4.4}$$

The threat score (TS), also called the critical success index (CSI), has the advantage that it takes into account that occurrences may be reflected less frequently in a dataset than non-occurrences (Wilks, 1995: 240). This is quite common in avalanche forecasting datasets. This score is a refinement of the hit rate and excludes correct non-avalanche forecasts. Even though the forecast errors (false stable/false unstable) contribute to the calculation of the score equally it does give a better indication of the quality of the forecast than the hit rate. The threat score is given by

$$TS = CSI = \frac{a}{a+b+c} \qquad \textit{Equation 4.5}$$

The worst possible threat score is zero, indicating that none of the avalanche days were correctly forecast.

The probability of detection (POD) is the fraction of those occasions when the forecast event occurred and was forecast (Wilks, 1995:240). It is calculated by

$$\text{POD} = \frac{a}{a+c} \qquad \textit{Equation 4.6}$$

A POD of 1 would mean that the likelihood that the event would be forecast, given that it occurred, is 100%.

The false-alarm rate (FAR) is computed as

$$\text{FAR} = \frac{b}{a+b} \qquad \textit{Equation 4.7}$$

and is the proportion of forecast avalanche days, when no avalanche occurred. An FAR score of 0 indicates that no false unstable forecasts were made. This score alone is not valuable in avalanche forecasting, because false unstable errors are considered less important than fast stable errors. However, in combination with other scores, FAR is of interest, because higher values would provide a reason to lose trust in the forecasting model.

The bias (B) is a comparison of the average forecast with the average observation given as

$$\text{B} = \frac{a+b}{a+c} \qquad \textit{Equation 4.8}$$

A bias of 1 indicates that the number of forecast avalanche days is the same as the number of avalanche occurrences in the dataset. However this does not mean that all avalanche days were correctly forecast. A bias greater than 1 indicates that the number of avalanche days forecast is greater than the number of avalanche days observed. A bias smaller than 1 is to be interpreted conversely. Again this score on its own does not give sufficient information on the quality of the forecast.

The Heidke and the Kuipers skill scores are similar. As mentioned earlier, skill scores compare a reference forecast to a set of forecasts. The difference between these two skill scores is in the denominator where the Kuipers skill score uses the hit rate for random forecasts that are constrained to be unbiased (Wilks, 1995:249). The Kuipers skill score has the advantage that a random and a constant forecast (e.g. only avalanche days forecast) receive a zero, indicating a random forecast. A skill score of 1 indicates a perfect forecast and a negative score a less accurate forecast than the random forecast. Also Wilks (1995: 250) mentions “the contribution made to the Kuipers score by a correct no or yes forecast increases as the event is more or less likely, respectively”. Because of the advantages of the Kuipers skill score it will be used for verification purposes in this thesis and is computed as

$$\mathrm{KSS} = \frac{(a+d)/n - ((a+b)(a+c)+(b+d)(c+d))/n^2}{1-((a+c)^2+(b+d)^2))/n^2}$$

where n = a + b+ c + d is the total number of forecasts.

In conclusion, most of these measures, excepting the hit rate, contribute to the verification process. However none of them is sufficient to describe the quality of an avalanche forecast on its own. Even in combination, the higher consequences of false stable forecasts compared to false unstable forecasts are not considered. Nevertheless the threat score (TS), the unweighted average accuracy (UAA), the probability of detection (POD), the false-alarm rate (FAR) and the Kuipers skill score (KSS) will be used in this thesis for forecast verifications.

5. Effect of normal load on the shear strength of persistent weak layers

5.1 Introduction

The effect of normal load on the strength of granular materials is well known from soil sciences, where the shear strength is determined by cohesive forces between particles and frictional resistance when particles slide over one another (Marshall and Holmes, 1979: 212). In snow research it is common to use the shear frame test as described in Section 4.2 to measure the strength of a layer. This requires that the snow above the measured layer is removed, thus reducing the normal load on the weak layer during the test (Jamieson and Johnston, 1998). In contrast to measuring the “internal friction” during slow shearing of soils, a normal load correction for rapid (brittle) shear testing of snow is required since the shear strength data for snow in this thesis and most field studies are in the brittle range.

Roch (1966) developed empirical formulas to adjust the strength measured with a shear frame to the normal load effect for different crystal forms. His equation for decomposed and fragmented crystals fitted a dataset from Canada well (Jamieson and Johnston, 1998), whereas his adjustment for depth hoar did not fit the surface hoar data, as might have been expected. While Roch (1966) reported an increase in shear strength for depth hoar, Jamieson and Johnston (1998) observed no significant increase in three out of five series and a decrease in the strength of surface hoar layers in two out of five series. The authors assumed that the normal load adjustment for persistent weak layers could be neglected considering their limited dataset. However, a normal load effect can be observed on layers of faceted crystals and surface hoar layers as Jamieson et al. (2001) noted that stability indices without a normal load adjustment were unrealistically low for deep weak layers. The aim of this chapter is to find an empirical model to determine the normal load effect on persistent weak layers using the five series from Jamieson and Johnston (1998) and an additional three series.

5.2 Field methods and dataset

Strength measurements were done with a shear frame, as explained in Section 4.2, in a level study site on horizontal layers. In addition to the usual testing of a layer, weights of 0.3, 1.0, 3.0, and 6.0 kg were placed on top of a 0.025 m^2 shear frame representing normal loads of 0.12, 0.39, 1.18 and 2.35 kPa respectively (Figure 5.1). Assuming a typical average slab density of 180 kg/m^3 these normal loads correspond to a slab thickness of 7, 22, 67 and 133 cm respectively. Routinely five to eight tests were performed on one layer with each weight. In a study conducted in the European Alps, Roch (1966) used a similar approach, except he used a 0.01 m^2 shear frame with normal loads were 1.0, 2.0, and 3.0 kg, thus the normal stresses he applied were larger due to the smaller shear frame size, corresponding to 0.98, 1.96 and 2.94 kPa or a slab thickness of approximately 56, 111 and 166 cm, respectively.

Figure 5.1: Photo of shear frame test with added weight and loaded distribution plate added digitally.

Subsequently the Daniels strength, the shear strength of an arbitrarily large failure area, was calculated using Föhn's (1987a) equations for shear frame size adjustments.

The Canadian dataset now consists of eight series measured between 1995 and 1998 in the Columbia and the Rocky Mountains. Six series were performed on surface hoar, one on rounded facets and one on decomposed and fragmented crystals and surface hoar, totalling 34 single measurements. Since the study of Jamieson and Johnston in 1998 the dataset has almost doubled.

In regard to the aim of finding a normal load effect for persistent weak layers (surface hoar, facets and depth hoar) two series on depth hoar (eight data points), measured by Roch in 1966, were included in the analysis, even though the observed strength increase with load was higher in his dataset. Unfortunately the dataset is still limited, but until other data become available on the normal load effect on the strength of persistent weak layers it is the best available.

5.3 Methods

The equation to determine the strength adjusted to the normal load is (Roch, 1966):

$$\Sigma_{ø} = \Sigma + \sigma_{zz} ø(\Sigma, \sigma_{zz}) \qquad \textit{Equation 5.1}$$

where z is measured from the snow surface normal to the slope, $ø(\Sigma, \sigma_{zz})$ is the normal load adjustment and the normal load is:

$$\sigma_{zz} = \rho g h \cos^2 \Psi \qquad \textit{Equation 5.2}$$

where ρ is the average slab density, g is the acceleration due to gravity, h is the slab thickness measured vertically and Ψ is the slope inclination, which is 0° in all the experiments used in this study.

Equation 5.1 is similar to the Coulomb-Mohr law for sliding friction between bodies from soil science (Hillel, 1980: 338), though Roch (1966) determined a non-constant relationship for ø as a function of the normal load and the shear strength dependent on the crystal type of the failure layer.

The normal load adjustment ($ø(\Sigma, \sigma_{zz})$) is the difference between the shear frame measurements with and without the weights. Föhn and Camponovo (1997) found that shear stress due to the slab increases linearly with slab depth. Simple regression analysis was used to find an empirical model to calculate the normal load adjustment. Additionally polynomial regression was tried, because it improved the fit of the regression model, indicating that for deeper layers the normal load adjustment is higher than for shallower layers. For both regression analyses no intercept was required in order to fit Equation 5.1.

5.4 Analysis

In most series there was an increase of strength with an increase of load. The greatest decrease measured was –0.062 kPa for a normal load of 0.39 kPa and the highest increase was 1.26 kPa for a normal load of 2.35 kPa including Roch's data. Excluding the depth hoar data the greatest decrease equaled –0.062 kPa and the highest increase 0.5 kPa indicating that Roch's data biases the strength calculation towards higher strength values. The median strength increase was 0.14 kPa (0.11 kPa excluding Roch's data), which means that half of the observations are between the minimum and 0.14 kPa (0.11 kPa). The dataset is unbalanced with more measurements where less normal load was applied. Unfortunately the strength adjustment is especially interesting for deeper layers, for which less data were available.

The polynomial regression analysis was rejected because the beta values of the coefficients in the regression equation were not significant.

A multiple regression analysis, using Σ and σ_{zz} as independent variables to determine the normal load adjustment, resulted in an insignificant beta value for the strength and an insignificant intercept.

In the simple regression analysis the following empirical equations for the normal load adjustment were found for the entire dataset (Equation 5.3a) and when excluding the depth hoar data from Roch (Equation 5.3b):

$$ø(\Sigma, \sigma_{zz}) = 0.32 \qquad \textit{Equation 5.3a}$$

with an $r^2 = 0.82$, $p = 10^{-16}$ and

$$ø(\Sigma, \sigma_{zz}) = 0.21 \qquad \textit{Equation 5.3b}$$

with an $r^2 = 0.77$, $p = 10^{-12}$

The analysis of the residuals for Equation (5.3a) showed normality at the 5% level but not at the 1% level (Shapiro-Wilk W = 0.939; p = 0.026) (Figure 5.2a), which fulfils the assumptions in this analysis. Normality of the residuals from Equation 5.3b was not rejected at both levels (Shapiro-Wilk W = 0.969; p = 0.434) (Figure 5.2b).

The constant normal load adjustments in Equations 5.3a and 5.3b are consistent with granular soils and in contrast to Roch (1966) who determined a non-constant relationship of ø as a function of Σ and σ_{zz}.

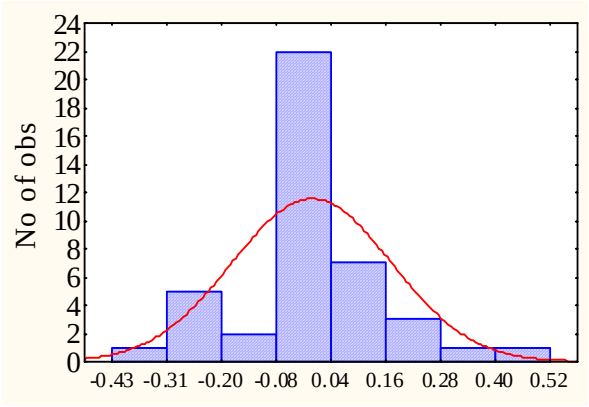

Figure 5.2a: Distribution of residuals in Equation 5.3a.

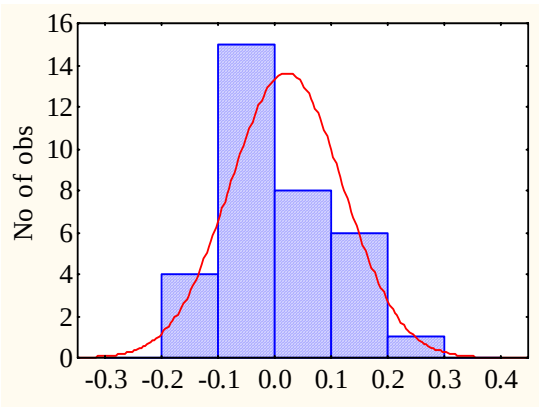

Figure 5.2b: Distribution of residuals in Equation 5.3b.

Even though some data points proved to be outliers, the model was not refined because the outliers were those with the highest normal loads applied. This error was expected because of the unbalanced dataset. However, it is crucial to predict the normal load for deeper weak layers and the heavier weights represented burial depth of 67-166 cm, depths important for destructive avalanches.

5.5 Model selection

In the following analysis of this thesis, Equation 5.3b was used to adjust the strength measurements for the normal load effect because of:

1) uncertainties in Roch's dataset, including grain types
2) possible overestimation of the strength and hence the stability using Roch's data, which could lead to false stable predictions with serious consequences in Canada
3) the preference of fitting the data from the transitional snow climate of the Columbia Mountains.

5.6 Strength adjusted to normal load

In Figure 5.3 the strength measurements and the calculated strength values (Equation 5.3b) for the Canadian dataset are plotted against the normal load. The strength calculations using the depth hoar equation from Roch (1966) are also included. Both models are quite similar, with correlations of $r^2 = 0.98$, however in most cases Equation 5.3b predicts better data points with higher normal loads compared to Roch's depth hoar equation.

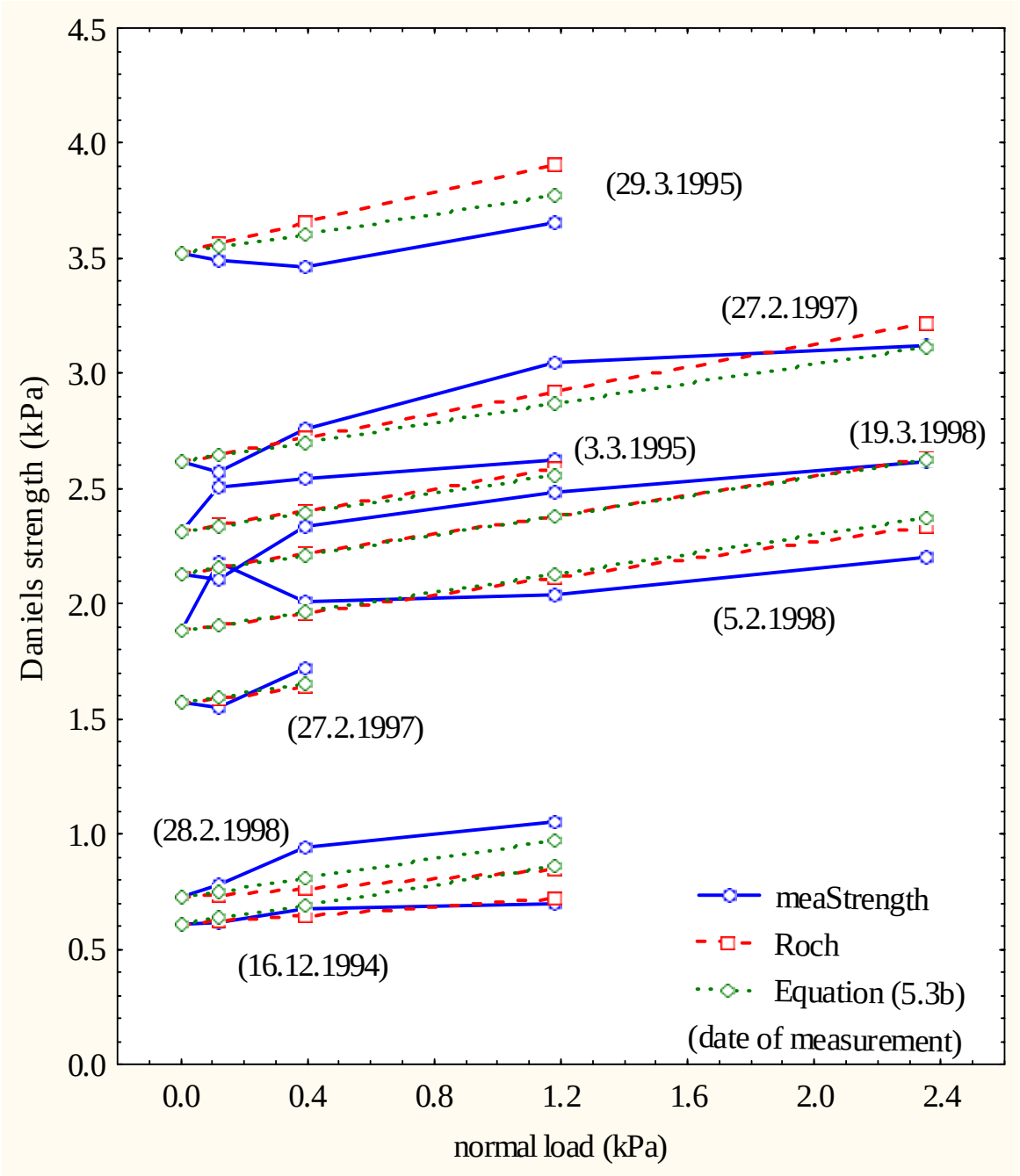

Figure 5.3: Measured and predicted effect of normal load on Daniels strength using Equation (5.3b) and Roch's depth hoar equation from 1966.

5.7 Summary

The brittle shear strength of persistent weak layers, as measured with the shear frame, showed a normal load effect. Even though the dataset used is limited, an empirical model was found to adjust the measured shear strength with a normal load effect. The model showed significant correlations and therefore in the following analysis the strength of persistent weak layers are adjusted to the normal load effect using Equation 5.3b.

6. Estimating the shear strength of persistent weak layers

6.1 Introduction

As previously stated, computer-assisted forecasting models have made little use of snowpack variables including stability indices, even though such indices have been shown to correlate with avalanche activity on a slope (Föhn, 1987b; Jamieson and Johnston, 1993) and regional scale (Jamieson, 1995: 125-137, 147-155). Stability indices are calculated using the shear strength of buried weak layers, which are potential failure layers for slab avalanches. However, shear strength measurements are time consuming and only performed on an intermittent basis at varying locations. In addition, only shear frame tests are easy to quantify, but are seldom used in forecasting operations. This is a drawback when considering the use of stability indices in daily avalanche forecasting models. Unless the shear strength can be extrapolated over time, stability indices are hard to incorporate into computer models.

Hägeli and McClung (2003) analyzed the characteristics of natural avalanche activity and list three types of persistent weak layers most relevant to avalanche occurrences in the Columbia Mountains of Canada: 1) faceted crystals, 2) surface hoar and 3) crust interfaces. They further found that weak layers of faceted crystals are often associated with significant avalanche cycles in the early season, and with sporadic, sometimes large releases throughout the entire season. Surface hoar layers in contrast, are generally prone to natural avalanche activity for about three to four weeks after burial (Hägeli and McClung, 2003).

Similar to the characteristics of natural avalanche activity, layers of surface hoar and faceted crystals form the failure layers for many skier-triggered avalanches in the Columbia Mountains. So far Chalmers (2001) has successfully estimated shear strength changes of surface hoar layers up to eight days after the day in which a manual snowprofile was observed. He considered in his analysis the shear strength of surface hoar layers for the first 30 days after burial; skier triggering appeared to be less likely after that time period. However in this thesis the slab thickness will be considered, because a layer older than 30

days but buried less than 100 cm might still be of concern for skier triggering. This might help to improve the forecast in the harder to predict transitional phases between unstable and stable.

Zeidler and Jamieson (2002) used a similar approach to forecast the shear strength of faceted layers. However, neither Chalmers (2001) nor Zeidler and Jamieson (2002) adjusted the measured shear strength for normal load in their analyses. In addition, more shear frame data on persistent weak layers are now available. This has enabled new regression analyses to be performed on bigger datasets using the shear strength adjustment for normal load as described in Chapter 5.

The specific objectives of this chapter are to:

1. explore the physical role of each predictor variable and their combined effects on the shear strength of faceted layers and surface hoar layers
2. develop models for estimating the shear strength of faceted layers and surface hoar layers (without shear frame tests) on a day in which a snowprofile was observed
3. extrapolate shear strength over time by developing models for the rate of change of shear strength of faceted layers and surface hoar layers, and
4. assess the combined models with shear strength measurements for faceted layers and surface hoar layers respectively.

In Section 6.3 the estimation of the shear strength of faceted layers will be discussed; Section 6.4 is concerned with the shear strength of surface hoar layers. As mentioned before, depth hoar layers are not common in the Columbia Mountains (Hägeli and McClung, 2003) and will not be considered.

6.2 Methods

6.2.1 Model building

The same equation used by Chalmers in 2001 in his Interval Model for estimating the shear strength of surface hoar layers was used in this analysis:

$$\Sigma_j^* = \Sigma_i^* + \Delta t_{ij}\ (\Delta\Sigma/\Delta t)_{ij}^* \qquad \textit{Equation 6.1}$$

where Σ_i^* and $(\Delta\Sigma/\Delta t)_{ij}^*$ are functions of snowpack observations on day *i*; Σ_i^* is the estimated shear strength on day *i* (kPa), $\Delta t_{ij} = t_j - t_i$ is the time interval between day *i* and day *j*, $(\Delta\Sigma/\Delta t)_{ij}^*$ is the estimated rate of change in shear strength (kPa d^{-1}) between day *i* and day *j*, and Σ_j^* is the forecast shear strength on day *j* (kPa).

Because Equation 6.1 is intended to estimate the shear strength of a persistent weak layer a number of days after the manual snowpack measurements, it is referred to as the Forecasting Model.

6.2.2 Response variables

The Forecasting Model required two formulas to be developed to describe:

1. the shear strength on the day of manual snowpack observations (Σ_i^*) and
2. the rate of change of shear strength between the day of the last snowpack observation and the day to be forecast $(\Delta\Sigma/\Delta t)_{ij}^*$.

6.2.3 Predictor variables

The set of predictor variables included snowpack and weather observations as well as elaborated variables (Table 6.1).

Table 6.1: Possible predictor variables; facets (FC), surface hoar (SH)			
Abbreviation	**Explanation**	**FC**	**SH**
Age	Age of the weak layer (days); number of days since buried	✓	✓
HS	Height of snowpack (cm)	✓	✓
Slope	Inclination of slope (°)	✓	✓
Load	Weight per unit area of overlying snow (kPa)	✓	✓
H	Thickness of overlying slab (cm)	✓	✓
HH	Hand hardness exponent (Geldsetzer and Jamieson, 2001)	✓	~
T+5	Temperature 5 cm above the weak layer (°C)	✓	✓
T-5	Temperature 5 cm below the weak layer (°C)	✓	✓
TG	Temperature gradient measured over 10 cm across failure plane of weak layer (°C/m)	✓	✓
Ta/HS	Average snowpack temperature gradient (°C/m)	✓	~
Twl	Temperature of the weak layer (°C)	✓	✓
Emin	Minimum grain size of weak layer crystal (mm)	✓	✓
Emax	Maximum grain size of weak layer crystal (mm)	✓	✓
Ta	Air temperature (°C)	✓	~
Thick	Thickness of the weak layer (cm)	✓	✓
SlabDens	Density of overlying slab (kg m^{-3})	✓	✓

In the analysis of the shear strength change rate, the initial shear strength of the weak layer was also included.

The predictor variables were chosen by their availability. The models developed in this thesis are intended to be valuable for practical use in forecasting operations. The predictor variables listed in Table 6.1 are mostly standard observations in a study plot; those that are not would be easy to measure additionally. Average values of the predictors, as considered in Zeidler and Jamieson (2002), to predict the shear strength change rate of faceted layers were excluded here because this would require the measurement of snowpack variables on a daily basis.

Each predictor variable is thought to have an influence on the shear strength and the rate of shear strength change of a persistent weak layer. However the snow temperature variables T+5, Twl and T-5 do have a similar physical meaning and are highly correlated (Pearson $r > 0.94$) in all datasets. T-5 correlated most strongly with the response variable shear strength ($r = 0.49$) compared to $r = 0.44$ and $r = 0.37$ for T+5 and Twl respectively for faceted layers. A similar pattern was observed for layers of surface hoar with a correlation of $r = 0.39$ of shear strength and T-5 compared to $r = 0.38$ for both T+5 and Twl.

Consequently T-5 was chosen for additional analysis. Additional intercorrelations of the predictor variables and exclusions from multiple regression analysis will be discussed in Section 6.3.6 and 6.4.6.

6.2.4 Analytical methods

Correlation analysis was applied to assess the relative importance of each individual predictor variable and to provide a basis for discussing the physical effect of the predictors on the response variables Σ_i* and $(\Delta\Sigma/\Delta t)_{ij}$*.

Simple linear regression analysis was used to examine the relationship (Allen, 1997: 16) between the most significant variables (5% level) from the correlation analysis and each of the response variables Σ_i* and $(\Delta\Sigma/\Delta t)_{ij}$*.

Multivariate analysis was used to model the relationship between the predictor variables and the two response variables Σ_i* and $(\Delta\Sigma/\Delta t)_{ij}$*.

The performances of the simple and the multivariate regression models were compared.

The regression analyses (simple and multivariate) were performed using the procedure:

1. initial regression,
2. refinement via outliers, where outliers were defined as any observation with a value of the standardized residuals of less than –3.0 or greater than 3.0 (Burt and Barber, 1996: 456),
3. analysis of identified outliers to assess whether they can be physically explained,
4. refined regression analysis, and if necessary refinement via assessment of outliers again,
5. residual analysis to determine variance and distribution, and
6. formulation of a final regression equation and assessment of goodness of fit.

As mentioned in the previous Section some of the potential predictor variables may be correlated with other predictors. This problem is known as multicollinearity (Mendenhall and Sincich, 1996: 355). Though it is quite common that a correlation between predictors exists in multiple regressions, serious multicollinearity can cause problems including high

standard errors of the beta values and consequently reduced reliability and misleading results (Mendenhall and Sincich, 1996: 355). The problem of multicollinearity is addressed in this thesis as follows:

1. correlation matrix for each of the datasets are presented before the multivariate regression analysis to identify potential problems regarding correlations. Farrer and Glauber (1967) suggest unacceptable collinearity when $r > 0.8$ or 0.9. In this thesis r-values greater than 0.8 are thought to be of concern.
2. exclusion of one of the cross-correlated variables from the analysis if one of the correlated variables does not add physically relevant information to the regression.
3. running stepwise regression as screening regression to select a good set of predictors from predictor variables that are almost always mutually correlated (Wilks, 1995: 188). In the backward stepwise regression all potential predictors are first considered and at each step the least important predictor variable is removed from the regression equation. The least important is the variable with the smallest coefficient in an absolute value sense with respect to its estimated standard error. The remaining variables are recomputed at each step because the predictors are mutually correlated (Wilks, 1995: 189). Medenhall and Sincich, 1996: 361) recommend stepwise regression as screening procedure to determine which of the correlated independent variables to drop.
4. calculation of the tolerance of the predictor variables included in the stepwise regression and rejection of regression when the tolerance of any predictor variable is lower than 0.2, indicating a problem with multicollinearity (Garson, n.d.). A tolerance of 0.1 indicates extreme multicollinearity (Mendenhall and Sincich, 1996: 357). Judd and McClelland (1989: 168) state that computer programs warn at a tolerance smaller than 0.001 or 0.01 because the predictor is too redundant to be added to the model. Considering this and that a tolerance of 0.1 indicates extreme multicollinearity it is reasonable to exclude predictors with a tolerance of less than 0.2 in this analysis. The tolerance is calculated as $1 - r^2$ where r^2 is the multiple coefficient of determination for the model that regresses each independent variable on the remaining independent variables in the regression.

6.3 Facets

6.3.1 Dataset

Data from the Columbia and the Rocky Mountains of western Canada were selected to develop a model to forecast the shear strength of faceted layers during the years 1993-2004. The selected layers of faceted crystals were

- November facets (NOV),
- near surface facets (NS), and
- facets on or below crusts in the upper snowpack (CRUST).

Each selected layer developed as a result of three main faceting processes (see Section 2.2).

Excluded from the analysis were:

- time series where the weak layers were less than 40 cm above the ground because of inconsistencies in the measurements most likely due to spatial variability,
- measurements with weak layer temperatures $> -1°C$, to avoid the late season effect where the strength of a weak layer (as measured with a shear frame) decreases, but does not result in increased avalanche activity,
- measurements where the slab thickness above the weak layers was less than 5 cm, because frame placements were difficult and might have affected the measured strength values,
- sets of measurements for which shear frame operators noted that the measurements were inconsistent or the fractures non-planar, such as observed on November facet layers in the Rocky Mountains, and
- one complete time series for model testing purposes (from an intermountain snow climate).

This reduced the dataset to 19 time series with a total of 102 shear strength measurements and 83 shear strength changes. The grains in 54% of the layers consisted of ICSSG (International classification of seasonal snow on the ground) Type 4c (rounding

facets), 18% of Type 4 (faceted crystals, not 4c), 16% of Type 4a (solid faceted crystals) and the remainder were classified as Type 4b (small faceted crystals), 3c (faceting rounds) or 4 as the second crystal form reported (Colbeck et al., 1990). Table 6.2 summarizes the time series with the location, number of measurements, and the formation of the facets. The BB-Lodge Plot is located in the Columbia Mountains, however the snowpack is shallower and classified as continental.

Table 6.2: Time series of faceted crystals. The time series for testing purposes is bolded and italic.

WL date	Location	N	Crust	NS	NOV	Continen-tal	Inter-mountain
5-Dec-93	BB-Lodge Plot	10	X			X	
15-Dec-93	BB-Lodge Plot	5	X			X	
22-Nov-96	Mt. St. Anne	13			X		X
29-Dec-96	Elk Study Slope	2		X			X
1-Mar-97	North Moose Log cut	2		X			X
1-Mar-97	Vermont Air Box	2		X		X	
1-Mar-97	Vermont Study Plot	2		X		X	
20-Nov-97	Mt. St. Anne	8			X		X
24-Dec-98	Mt. St. Anne	5		X			X
18-Nov-99	Schuss Lake	11			X		X
9-Feb-00	Bow Summit North	7		X		X	
26-Mar-01	Mt. St. Anne	5	X				X
16-Nov-01	Fidelity Study Slope	7			X		X
7-Jan-02	***Mt. St. Anne***	***12***	***X***				***X***
5-Apr-02	Mt. St. Anne	3	X				X
15-Jan-04	Flaming Corner MSA	4	X				X
15-Jan-04	Thunder Log Cut	3	X				X
3-Mar-04	Fidelity Study Slope	4	X				X
9-Mar-04	Mt. St. Anne	3	X				X
12-Mar-04	Fidelity Study Slope	6	X				X

6.3.2 Descriptive statistics and distribution of the response variables

The shear strength (Σ_i) in the dataset (N = 102) ranged from 0.23 to 8.46 kPa with a mean value of 2.58 kPa, although 75% of the observations had shear strength values of less than 3.18 kPa. As seen in Figure 6.1a the distribution of shear strength is highly positively

skewed and truncated at zero. This is expected because the shear strength cannot be negative and because there is a bias in the measurements towards more observations when the shear strength of the weak layer is lower. While in the first two weeks after burial of a weak layer the shear strength was generally measured 2-3 times per week, afterwards the measurement intervals become longer when the facets become stronger and the shear strength change slower. The hypothesis of normality for the shear strength distribution is rejected with the Kolmogorov-Smirnov test ($D = 0.216$, $p < 0.01$) and the Lilliefors test ($p < 0.01$).

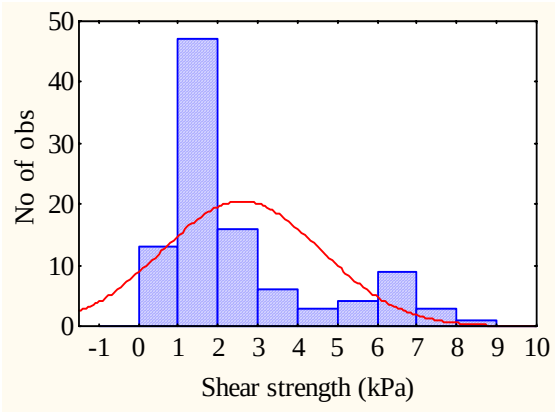

Figure 6.1a: Distribution of shear strength

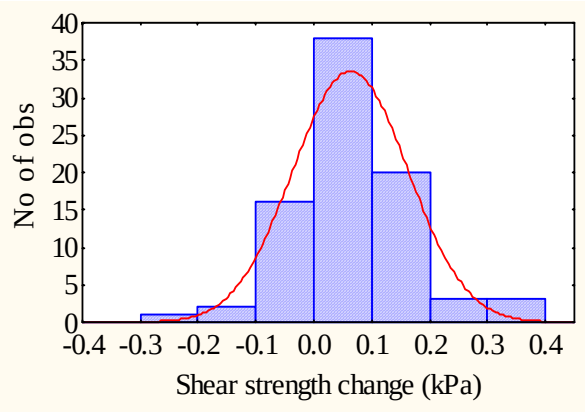

Figure 6.1b: Distribution of shear strength change rate

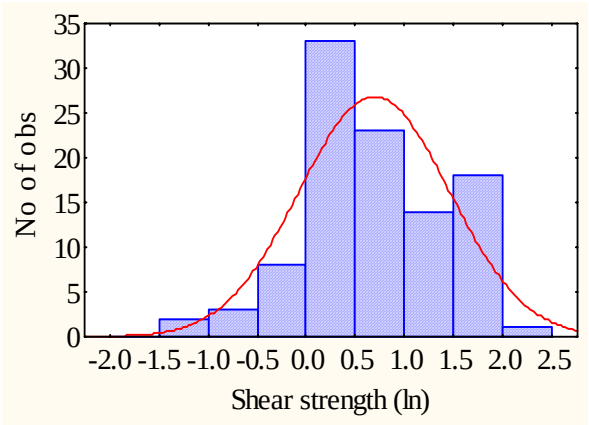

Figure 6.1c: Distribution of shear strength (logarithmic)

A logarithmic transformation of shear strength results in a distribution as seen in Figure 6.1c. The Kolmogorov-Smirnov test (D = 0.087, p <n.s.) and the Lilliefors test (p < 0.1) do not reject the assumption of normality.

The distribution of the shear strength change rate (Figure 6.1b) shows that shear strength gains are more common than shear strength losses over time with a minimum measured shear strength change rate of –0.28 kPa and a maximum gain in shear strength of 0.36 kPa per day. The mean and the median were both approximately 0.06 kPa, indicating that 50% of the measurements gained less than 0.06 kPa per day or diminished in shear strength and 50% gained more than 0.06 kPa. The hypothesis of normality for the shear strength change rate was not rejected with the Kolmogorov-Smirnov test (D = 0.105, p < n.s.), though the Lilliefors test (p < 0.05) only gives an indication that the shear strength change rate is from a normal distribution at the 5% level of significance and not at the 1% level. The Lilliefors test has the advantage over the Kolmogorov-Smirnov test in that the test accounts for the mean and standard deviation being estimated from the data being fitted. However, normality is accepted in this case.

6.3.3 Exploring the dataset

The objective of this thesis is to improve the ability to forecast skier-triggered avalanches in the Columbia Mountains of Canada. For this to be achieved, careful selection of the most accurate datasets is required. In this analysis, shear strength measurements from the Columbia Mountains (intermountain snow climate) and the Rocky Mountains (continental snow climate) were used. The question is whether a better prediction model can be found by using a larger dataset, including data from a continental and an intermountain snowpack, or by using a dataset solely consisting of data from the intermountain snowpack. The growth of facets strongly depends on the temperature and temperature gradient of the snow (see Section 2.2). The continental snowpack is often thinner and the temperatures colder, which might provoke a different shear strength behavior over the course of the winter. In Figure 6.2 the shear strength to load ratio, (and load will later prove to be one of the most important predictor variables), is shown graphically to solely access whether the continental and the intermountain data can be used

in conjunction without interpreting the data itself. Consequently a regression line is not included in the graph. The distribution of the continental data seems to fit in with the overall distribution of the entire dataset, possibly because November facets from the Rocky Mountains were excluded prior to the analysis. The November facets are exposed to stronger temperature gradients over much of the winter and the resulting weakening effects of the faceting process compete with the strengthening effects of densification and pressure sintering. Except for the exclusion of the November facets a further distinction between continental and intermountain datasets will not be made.

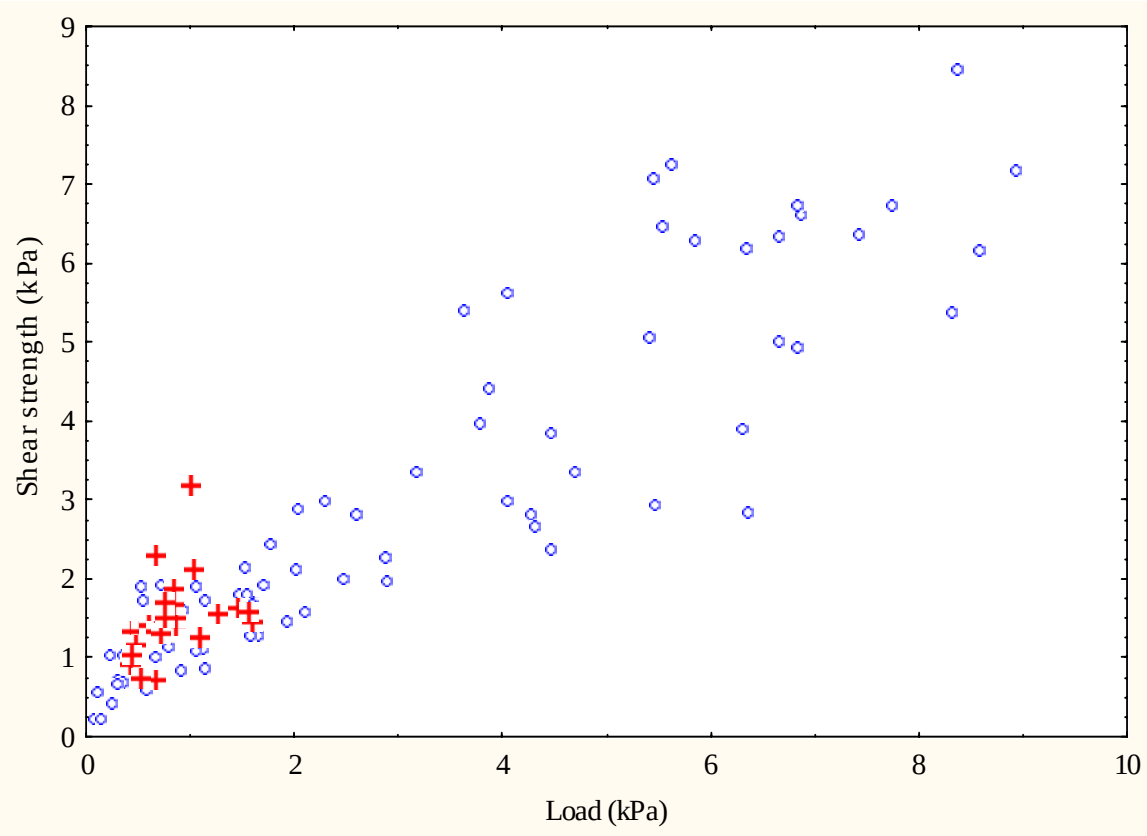

Figure 6.2: Scatterplot of shear strength vs Load. Crosses represent data from a continental snowpack. Circles represent data from the intermountain ranges.

The burial depth of a weak layer is considered important for skier triggering, though deep slab avalanches may occur and surprise skiers. Therefore, it is interesting to compare the shear strength of deep and shallow weak layers. In Figure 6.3 the shear strength is again plotted against Load but in this case the data points are highlighted when the depth of the weak layers exceeded 170 cm. Even though it is said that a skier is unlikely to trigger an avalanche below a weak layer depth of 100 cm, 170 cm was chosen because, as seen in Figure 6.3, the variance in shear strength seems to increase with Load (about 3 kPa) and depth (about 170 cm).

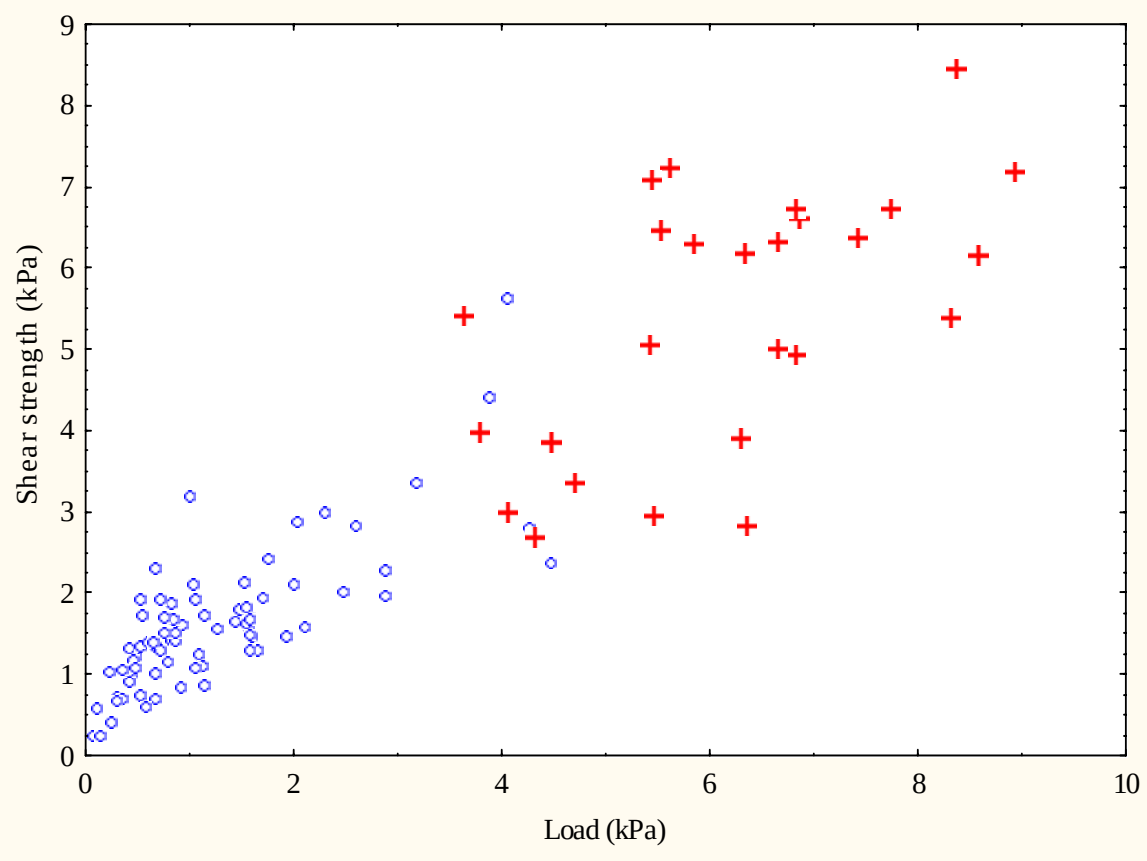

Figure 6.3: Scatterplot of load vs shear strength. Crosses mark the data points measured at a greater depth than 170 cm.

The shear strength variance is greater at larger depths than at shallower depths. As a result, two different datasets will be considered:

1) combined: entire dataset (N = 102)
2) H<170: data points where the weak layers were not deeper than 170 cm (N = 76)

Deeper burials alone were not considered because of the limited dataset, even though this might improve the forecast for deep slab avalanches.

A distinction into datasets only considering near surface facets, facets on or below crusts or November facets showed no advantage, because the November facets, which exhibit greater shear strength due to deeper burials, are mostly excluded in the H<170 dataset.

The near surface facets and the facets on and below crusts seem to have a similar load/shear strength relationship even though the near surface facets are generally stronger (Figure 6.4).

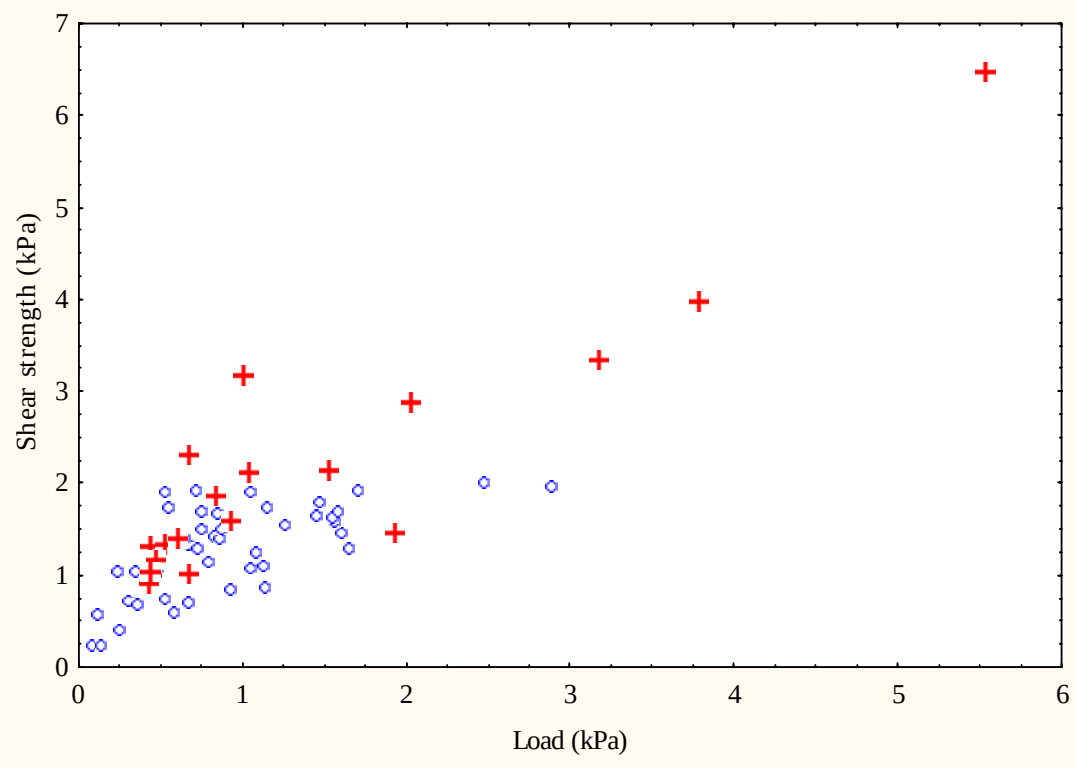

Figure 6.4: Scatterplot load vs shear strength. Crosses mark near surface facets.

6.3.4 Correlation analysis

The non-normal distribution of the response variable shear strength and the weak indication of a normal distribution of the shear strength change rate as shown in Section 6.3.2 suggest the use of non-parametric correlation analysis, such as the Spearman rank correlation. In addition the relationship between the response variables and the predictor variables are likely to be monotonic increasing or decreasing but not necessarily linear and therefore rank correlation analysis is appropriate (Burt and Barber, 1996: 394).

In Table 6.3 the Spearman rank correlations with shear strength are shown for the combined dataset.

Table 6.3. Spearman rank correlations with shear strength for the combined dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
Load	102	0.882	<1E-17
SlabDens	102	0.858	<1E-17
H	102	0.839	<1E-17
HH	99	0.741	<1E-17
Age	102	0.723	1.0E-17
T-5	87	0.527	1.6E-7
HS	102	0.480	3.3E-7
Emax	101	0.384	7.4E-5
Ta	99	-0.296	0.003
Emin	101	0.254	0.010
Slope	102	-0.234	0.018
Thick	*97*	*-0.065*	*0.529*
TG	*87*	*-0.044*	*0.686*
Ta/HS	*99*	*-0.017*	*0.868*

Most predictor variables were significantly correlated ($p < 0.05$) with shear strength, except Thick, TG and Ta/HS. Even though five additional years of data were available since Johnson (2000) analyzed a similar dataset, due to the exclusions mentioned in Section 6.3.1, the dataset included fewer shear strength measurements. In contrast to Johnson (2000: 57-63) the correlation with temperature gradient was not significant in Table 6.3. A possible explanation is that since 2000 facets on crusts close to the snow surface were tested more frequently where the relevant temperature gradient within millimetres of the

crust was difficult to measure. Another difference is that in this analysis, grain size Emax was found to be significant with a positive correlation. This could be due to the continued testing of layers of large crystals after they became relatively strong.

Using only data points where the slab thickness did not exceed 170 cm, the correlations are shown in Table 6.4.

Table 6.4. Spearman rank correlations with shear strength for the H<170 dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
SlabDens	76	0.761	1.6E-15
Load	76	0.752	5.2E-15
H	76	0.643	3.7E-10
HH	76	0.571	7.1E-8
Age	76	0.512	2.3E-6
T-5	61	0.290	0.024
Emin	*75*	*0.214*	*0.066*
HS	*76*	*0.209*	*0.070*
Emax	*75*	*0.210*	*0.071*
Slope	*76*	*-0.206*	*0.074*
Ta	*74*	*-0.180*	*0.125*
TG	*61*	*-0.109*	*0.402*
Thick	*73*	*-0.042*	*0.723*
Ta/HS	*74*	*-0.034*	*0.774*

The slab density, load, slab thickness, hand hardness and age were the most significant factors; T–5 was found to be significant at the 5% level but not at the 1% level. All other variables showed no significant correlation.

The highest ranked predictor variables will be used in the simple regression analysis in Section 6.3.5 and all significant predictors in the multivariate regression in Section 6.3.6.

The Spearman rank correlations between each predictor and the shear strength change rate showed only a weak correlation with Emax ($p = 0.01$) for the combined dataset (Table 6.5) and no significant correlations for the H<170 dataset (Tables 6.6).

Table 6.5. Spearman rank correlations with shear strength change rate for combined dataset. Insignificant variables are in italics (p > 0.05)

Variable	Valid N	Spearman R	p-level
Emax	82	-0.275	0.013
Slope	*83*	*-0.199*	*0.072*
TG	*70*	*0.209*	*0.082*
Age	*83*	*-0.187*	*0.090*
Strength	*83*	*-0.180*	*0.103*
Emin	*82*	*-0.159*	*0.154*
HS	*83*	*0.129*	*0.247*
Twl	*83*	*-0.126*	*0.256*
SlabDens	*83*	*-0.121*	*0.278*
Load	*83*	*-0.110*	*0.323*
Thick	*80*	*-0.106*	*0.351*
H	*83*	*-0.093*	*0.401*
T-5	*70*	*-0.070*	*0.564*
Ta	*80*	*-0.058*	*0.611*
Ta/HS	*80*	*-0.027*	*0.814*
HH	*80*	*-0.023*	*0.837*
T+5	*70*	*-0.019*	*0.879*

Table 6.6. Spearman rank correlations with shear strength change for H<170. Insignificant variables are in italics (p > 0.05)

Variable	Valid N	Spearman R	p-level
Age	*62*	*-0.246*	*0.054*
Strength	*62*	*-0.238*	*0.063*
Emax	*61*	*-0.228*	*0.077*
Thick	*60*	*-0.223*	*0.086*
HS	*62*	*0.209*	*0.103*
Slope	*62*	*-0.198*	*0.122*
TG	*49*	*0.198*	*0.172*
SlabDens	*62*	*-0.155*	*0.230*
Load	*62*	*-0.154*	*0.232*
Emin	*61*	*-0.149*	*0.251*
H	*62*	*-0.135*	*0.296*
T+5	*49*	*0.151*	*0.300*
HH	*62*	*-0.087*	*0.501*
T-5	*49*	*0.090*	*0.540*
Ta/HS	*60*	*0.032*	*0.811*
Ta	*60*	*0.028*	*0.833*
Twl	*62*	*-0.026*	*0.839*

Consequently, the regression analysis was omitted for shear strength change rate and instead the long term average loading rates for the two snow climates as well as daily loading rates were used to predict the change in shear strength between snowpack observations (Section 6.3.7).

Four of the top-ranked predictors that yielded significant Spearman rank correlations with shear strength in both datasets were of particular interest for the regression analysis in the next section:

- **Load:** The positive rank correlation showed that greater loads typically overlie stronger facet layers. Load causes densification (Kojima, 1967; Conway and Wilbour, 1999) and pressure sintering between the load-bearing crystals and increased bonding.
- ***H:*** The positive rank correlation indicates that thicker slabs typically overlie stronger facet layers. This is expected since slab thickness is strongly correlated to load.
- **SlabDens:** The positive rank correlation indicates that denser slabs typically overlie stronger facet layers. Because denser slabs are usually older and apply more load than less dense slabs, the underlying facet layer is probably stronger due to densification and pressure sintering (in a snow climate that favours equilibrium metamorphism).
- **HH:** The positive rank correlation indicates that harder facet layers usually have greater shear strength. This is expected because hand hardness and shear strength are both measures of bonding. While hand hardness of thick layers is usually easier to measure than shear strength, hand hardness is partly subjective and is difficult to estimate for thin layers.

For a physical interpretation of other predictors, see Johnson (2000: 57-63).

6.3.5 Simple regression analysis

A simple linear regression of measured shear strength on the highest ranked predictor variables for the combined dataset and H<170 resulted in best r^2 values of 0.85 (Table 6.7) and 0.6 (Table 6.8) respectively using Load as a predictor variable and after removing the outliers as described in Section 6.2.2. The coefficient of variation (CV) is a standardized measure of the goodness of fit and allows the comparison of two or more variables with different means and standard deviations and is calculated by dividing the standard error of the estimate by the mean of the observed values and multiplied by 100 to express a percentage (Burt and Barber, 1996: 447).

In the combined dataset there was only one outlier, which exhibited an unrealistic shear strength loss of 1.07 kPa over the eight days since the last shear frame measurements possibly due to spatial variability. Two outliers were removed from the H<170 dataset. The first outlier was measured at a depth of 165 cm and had the highest measured strength value in the dataset of 5.6 kPa. The second outlier was measured at Bow Summit in a continental snowpack, where the weak layer was only buried 48 cm at an age of 47 days and the shear strength of 3.2 kPa was high, compared to measurements at that depth in the Columbia Mountains.

Table 6.7: Simple linear regression of shear strength on Load. Combined dataset.

N = 101	B	Standard Error B	Significance Level *p*	r^2	Standard Error of Estimate	Overall *p*	CV
Intercept	0.703	0.111	7.1E-9	0.85	0.780	< 1.0E-17	30
Load	0.760	0.032	< 1.0E-17				

Table 6.8: Simple linear regression of shear strength on Load. H<170 dataset.

N = 74	B	Standard Error B	Significance Level *p*	r^2	Standard Error of Estimate	Overall *p*	CV
Intercept	0.823	0.087	2.9E-14	0.6	0.457	4.5E-16	30
Load	0.594	0.057	4.5E-16				

The fit is better for the combined model, which includes deeper burials, possibly because of the larger dataset, and load seems to play a stronger role for more deeply buried layers. This is in accordance with Keeler (1969) who postulated that overburden pressure is most important below 180 cm. Slab density and slab thickness yielded less significant correlations; however, they are directly related to load. Even though the Spearman rank correlation yielded SlabDens with the highest correlation for the H<170 dataset, the regression was less significant (r^2 = 0.44, p = 3.8E-11) than with load (r^2 = 0.6, p = 4.5E-16). This result does not contradict the Spearman rank correlation, because regression is a parametric analysis. However, the aim of the regression analysis is to formulate a model to predict the shear strength and a suitable non-parametric prediction scheme is not available. Even though the assumption of normality of the response variable is violated, regression analysis can be used as long as the residuals are normally distributed and the residuals are approximately constant (homoscedastic). A plot of the residuals versus the predicted values (Figure 6.5a and 6.6a) shows that the variance increased with shear strength in both models, though not as apparent in the H<170 dataset, which could be expected, because higher loads (with higher variances) were excluded. The normal distribution of the residuals is not rejected for the combined dataset with the Kolmogorov-Smirnov test (D = 0.100, p < n.s.), but was with the Lilliefors p < 0.05) (Figure 6.5b). The normality of the residuals is not rejected for the H<170 dataset (Kolmogorov-Smirnov D = 0.081, p < n.s. and Lilliefors p < 1) (Figure 6.6 b). The negative residuals for the lowest predicted values illustrate that the regression overestimates shear strength for low values of load.

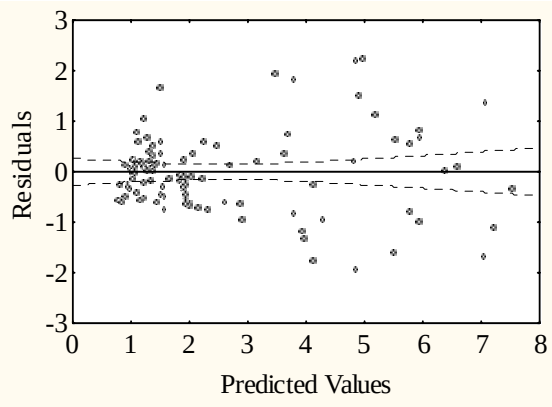

Figure 6.5 a: Scatter of residuals from the regression of shear strength on Load. Combined dataset.

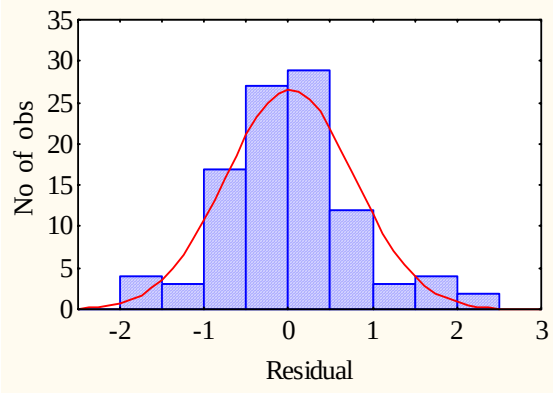

Figure 6.5 b: Distribution of residuals from the regression of shear strength on Load. Combined dataset.

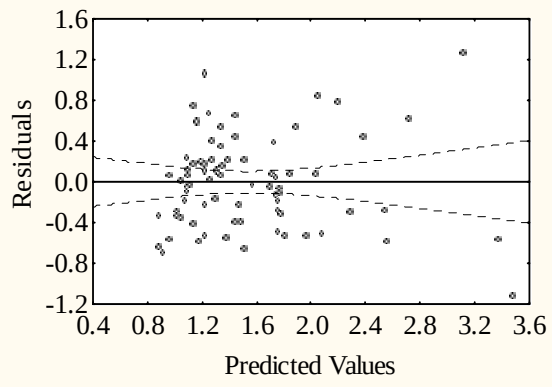

Figure 6.6 a: Scatter of residuals from the the regression of shear strength on Load. H<170 dataset.

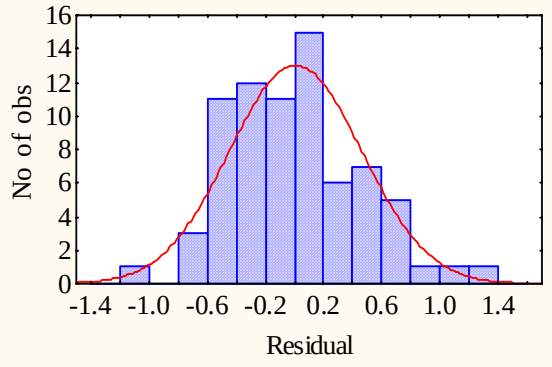

Figure 6.6 b: Distribution of residuals from the regression of shear strength on Load. H<170 dataset.

Figures 6.7a and b show the relationship between shear strength and load for both datasets. The higher variance for higher strength values is apparent. As shown in Section 6.3.2 a logarithmic transformation of shear strength results in a normal distribution of the response variable, however Figures 6.7c and d reveal that there is still a non-linear trend in the data. Additionally applying a logarithmic transformation on load results in a linear relationship as shown in Figures 6.7e and f.

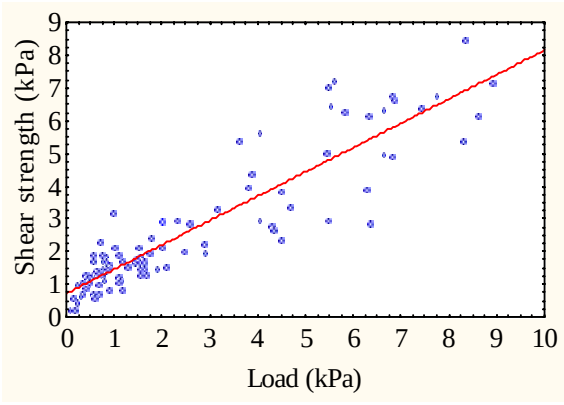

Figure 6.7 a: Shear strength vs Load. Combined dataset.

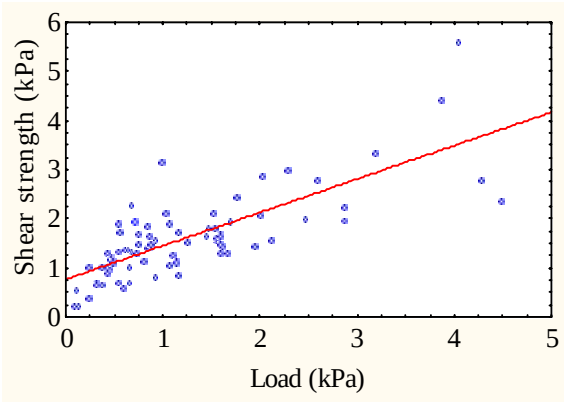

Figure 6.7 b: Shear strength vs Load. H<170 dataset.

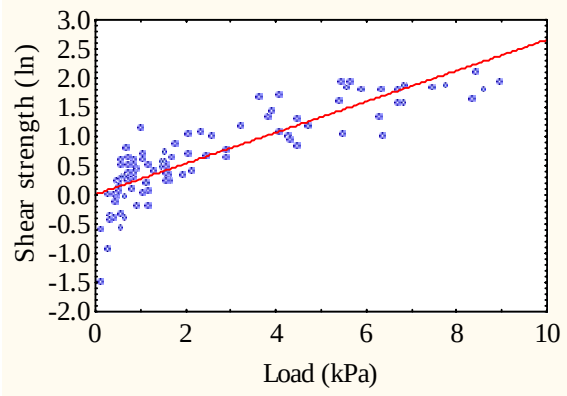

Figure 6.7 c: Shear strength (ln) vs Load. Combined dataset.

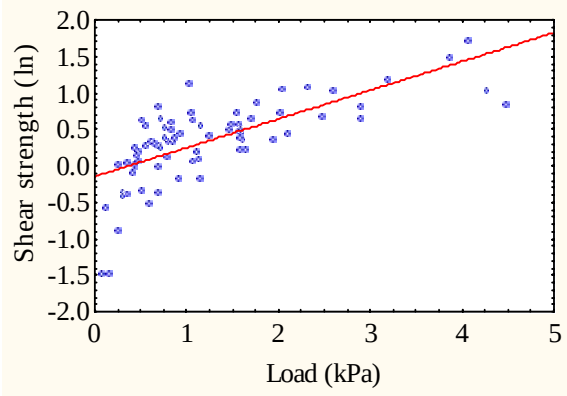

Figure 6.7 d: Shear strength (ln) vs Load. H<170 dataset.

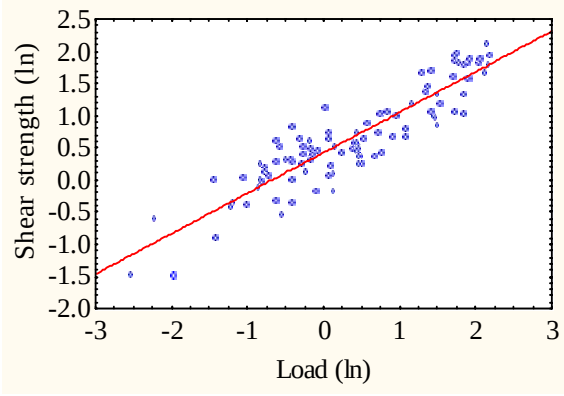

Figure 6.7 e: Shear strength (ln) vs Load (ln). Combined dataset.

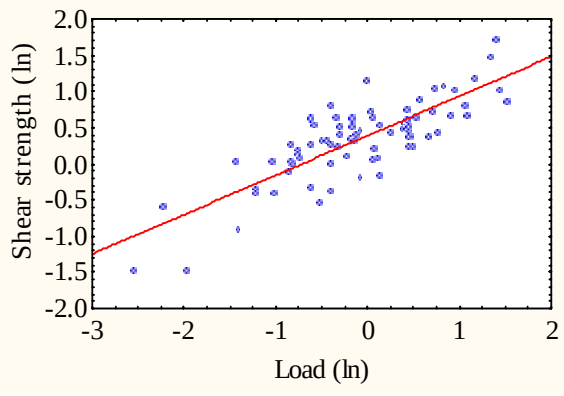

Figure 6.7 f: Shear strength (ln) vs Load (ln). H<170 dataset.

Consequently to further enhance the regression model, a logarithmic regression was considered to regress ln Σ on ln Load for both models to reduce the hetereoscedasticity as observed in the initial regression (Wilks, 1995: 172).

$$\ln \Sigma = \ln A + B \ln \text{Load} \qquad \text{(Equation 6.2a)}$$

which can be re-written as a power law

$$\Sigma = A\,\text{Load}^{B} \qquad \text{(Equation 6.2b)}$$

The variance of the residuals in the logarithmic transformation was stabilized (Figure 6.8a and 6.9a). Once again the residuals in both datasets (Figure 6.8b and 6.9b) showed to be from a normal distribution (Kolmogorov-Smirnov D = 0.058, p < n.s. and Lilliefors p < 1 (combined dataset) and Kolmogorov-Smirnov D =0.062, p < n.s. and Lilliefors p < 1 (H<170 dataset)).

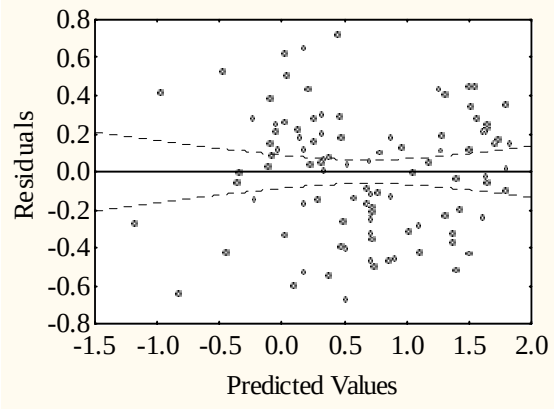

Figure 6.8a: Scatter of residuals from the regression of ln shear strength on ln Load. Combined dataset.

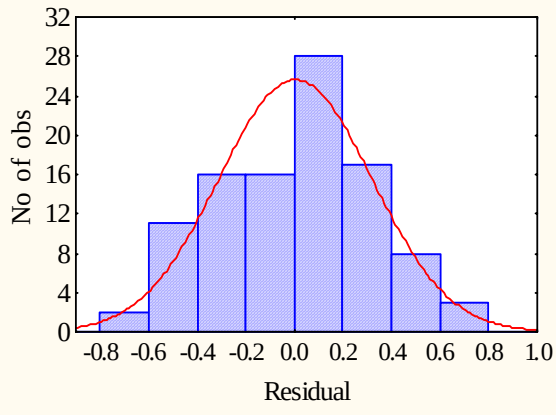

Figure 6.8b: Distribution of residuals from the regression of ln shear strength on ln Load. Combined dataset.

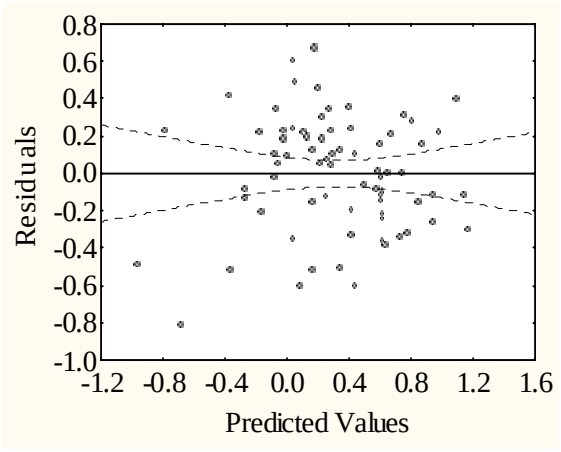

Figure 6.9a: Scatter of residuals from the regression of ln shear strength on ln Load H<170 dataset.

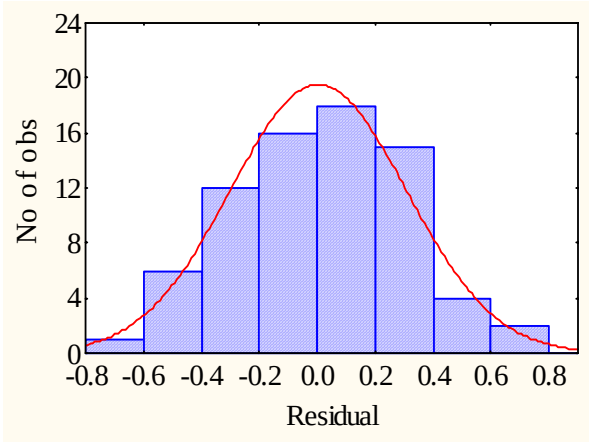

Figure 6.9b: Distribution of residuals from the regression of ln shear strength on ln Load H<170 dataset.

For the logarithmic regression, results for the combined dataset and H<170 cm are shown in Tables 6.9 and 6.10. The r^2 = 0.83 for the combined model was similar to the regression model before the transformation, whereas for the H<170 dataset the r^2 improved from 0.60 to 0.67. The coefficients in both models are similar. The calculation of CV in a logarithmic transformation is not meaningful in this dataset, because the mean is close to zero.

Table 6.9: Simple logarithmic regression. Combined dataset.

N = 101	B	Standard Error B	Significance Level *p*	r^2	Standard Error of Estimate	*p*
Intercept	0.427	0.033	<1.0E-17	0.83	0.316	<1.0E-17
ln Load	0.636	0.029	<1.0E-17			

Table 6.10: Simple logarithmic regression. H<170 dataset.						
N = 74	B	Standard Error B	Significance Level *p*	r^2	Standard Error of Estimate	*p*
Intercept	0.369	0.036	7.2E-16	0.67	0.305	<1.0E-17
ln Load	0.527	0.044	<1.0E-17			

The fit of the combined dataset to Equation 6.2a was substantially better (r^2 = 0.83) than that for H<170 (r^2 = 0.67), indicating that the predictive potential of load on the shear strength of facets is higher when the deeper burials are included. However, a better fit was not achieved when regressing shear strength on predictor variables other than load.

Re-writing the coefficients of Table 6.9 for the combined dataset and 6.10 for the H<170 dataset as Power Law models results in:

$$\Sigma = 1.53\ \text{Load}^{0.64} \qquad \text{(Equation 6.3a)}$$

$$\Sigma = 1.45\ \text{Load}^{0.53} \qquad \text{(Equation 6.3b)}$$

6.3.6 Multivariate regression

Table 6.11a and b show the correlation matrix of the predictors for both datasets. Strong correlations between predictors (r > 0.8) are highlighted.

Emax was excluded from further analysis because the correlation with Emin is almost at the critical correlation of 0.8 and measurements of Emax in the earlier years of the observations were not reported consistently. The correlation of H and Load is of most concern with the strongest correlation in both datasets. However both are included in the stepwise regression. The physical reasoning behind this is that H can offer additional information because thinner slabs are easier to trigger than thicker slabs with a similar load. Age is correlated with SlabDens, H and Load in the entire dataset; however, not in the H<170 cm dataset. The stepwise regression is run with and without age.

Table 6.11a: Correlation between predictors (r). Combined dataset. Facets.														
Variab	**Age**	**Slope**	**Emin**	**Emax**	**Thick**	**T-5**	**Ta**	**HS**	**H**	**SlabD**	**HH**	**TG**	**Ta/HS**	**Load**
Age	1													
Slope	-0.21	1												
Emin	0.49	-0.32	1											
Emax	0.48	-0.08	0.77	1										
Thick	0.16	-0.05	0.24	0.07	1									
T-5	0.56	-0.11	0.26	0.23	0.10	1								
Ta	-0.21	-0.01	-0.05	-0.03	-0.18	-0.11	1							
HS	0.14	0.10	-0.18	-0.01	-0.31	0.22	0.03	1						
H	0.73	-0.10	0.18	0.28	-0.02	0.47	-0.35	0.62	1					
SlabD	**0.80**	-0.34	0.35	0.35	-0.01	0.49	-0.14	0.31	0.76	1				
HH	0.53	-0.09	0.13	0.16	-0.13	0.28	-0.30	0.39	0.69	0.58	1			
TG	-0.01	0.05	-0.17	-0.07	-0.10	0.18	0.23	0.54	0.07	0.05	0.01	1		
Ta/HS	-0.03	0.01	-0.05	0.07	-0.28	0.04	0.57	0.48	0.13	0.05	0.03	0.46	1	
Load	**0.81**	-0.20	0.23	0.31	-0.05	0.50	-0.26	0.56	**0.96**	**0.86**	0.69	0.07	0.13	1

Table 6.11b: Correlation between predictors (r). H<170 dataset. Facets.														
Variab	**Age**	**Slope**	**Emin**	**Emax**	**Thick**	**T-5**	**Ta**	**HS**	**H**	**SlabD**	**HH**	**TG**	**Ta/HS**	**Load**
Age	1													
Slope	-0.17	1												
Emin	0.64	-0.26	1											
Emax	0.49	0.01	0.78	1										
Thick	0.44	-0.10	0.28	0.12	1									
T-5	0.41	-0.07	0.18	0.08	0.13	1								
Ta	-0.18	0.03	-0.11	-0.04	-0.23	-0.03	1							
HS	-0.47	0.25	-0.31	-0.15	-0.28	0.02	0.22	1						
H	0.26	0.12	0.12	0.18	0.22	0.20	-0.33	0.42	1					
SlabD	0.54	-0.28	0.31	0.23	0.14	0.29	-0.09	-0.08	0.43	1				
HH	0.18	0.08	0.12	0.09	-0.03	-0.05	-0.37	0.09	0.44	0.30	1			
TG	-0.16	0.06	-0.20	-0.08	-0.10	0.18	0.30	0.60	0.05	0.01	-0.06	1		
Ta/HS	-0.23	0.06	-0.09	0.03	-0.28	0.04	0.63	0.51	0.08	-0.09	-0.09	0.50	1	
Load	0.35	-0.02	0.18	0.21	-0.20	0.25	-0.30	0.34	**0.95**	0.63	0.45	0.04	0.06	1

The combined dataset for the *multivariate regression* for estimating shear strength included 82 shear strength measurements and the H<170 dataset included 64. The lower number of measurements available for the multivariate analysis compared to the univariate analysis reflects missing values in several of the predictor variables. In the analysis only the significant variables from the Spearman rank correlations in Table 6.3 and 6.4 were used. Stepwise regression (backwards) selected the predictors: Age, Slope and H in the combined dataset (Table 6.12) after two outliers that exhibited high strength values were removed because the observed faceted layers were only reported with thicknesses of 0.1 and 0.2 cm which made it difficult to perform shear frame tests. In addition, for these outliers, inconsistent frame placements were reported in the field book, though the fracture was planar, lending further support for the removal of these values from the analysis. The coefficient of determination (Table 6.12) indicates a better fit than in the univariate logarithmic regression (0.85 compared to 0.83).

Table 6.12: Multivariate backwards stepwise regression of shear strength. Combined dataset.

N = 80	B	Standard Error B	Significance Level *p*	Tolerance	Adj. r^2	Standard error of estimate	*p*	CV
Intercept	0.605	0.169	5.9E-4		0.85	0.760	< 1.0E-17	30
Age	0.023	0.006	7.9E-5	0.24				
Slope	-0.024	0.007	4.8E-4	0.94				
H	0.014	0.002	2.9E-8	0.25				

The variance of the residuals (Figure 6.10a) was again higher when stronger faceted layers were tested. The residuals were tested for normality which was not rejected with the Kolmogorov-Smirnov D = 0.08, p < n.s. and a Lilliefors p < 0.15 (Figure 6.10b).

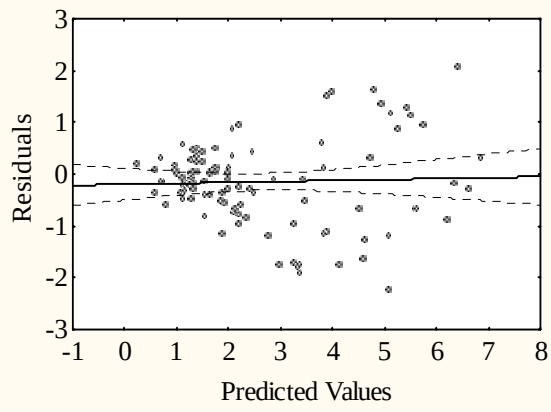

Figure 6.10a: Scatter of residuals against predicted values Multivariate regression. Combined dataset.

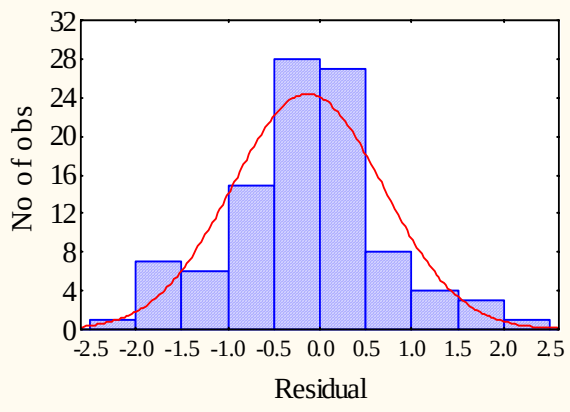

Figure 6.10b: Distribution of residuals. Multivariate regression. Combined dataset.

Although Age can be considered as a predictor variable, the interpretation is somewhat difficult. Consequently multivariate regressions were performed without using Age as a predictor. In the combined dataset, only Load was picked as a predictor variable by the backward multiple regression model and therefore the simple regression in Equation 6.3a applies.

In the multiple backwards regression analysis with the H<170 dataset the variables H and SlabDens were selected as predictor variables (Table 6.13). Two measurements were deleted as statistical outliers. Again the weak layer for these measurements was very thin (0.2 cm) and the field workers reported that the frame placement was difficult which resulted in high strength values. However the fit, compared to the simple linear regression improved from an r^2 of 0.67 to 0.71, but the intercept was not significant (Table 6.13).

Table 6.13: Multivariate backwards stepwise regression. H<170 dataset

N = 61	B	Standard Error B	Significance Level *p*	Tolerance	Adj. r^2	Standard Error of Estimate	*p*	CV
Intercept	-0.303	0.186	0.110		0.71	0.383	1.5E-15	24
H	0.007	0.001	3.0E-6	0.74				
SlabDens	0.008	0.001	1.2E-8	0.74				

Other than for the multivariate regression for the combined dataset the residuals are homoscedastic (Figure 6.11b). However the normality of the residuals is rejected with the Lilliefors $p < 0.05$, though not with the Kolmogorov-Smirnov test ($D = 0.115$, $p <$ n.s.) (Figure 6.11b).

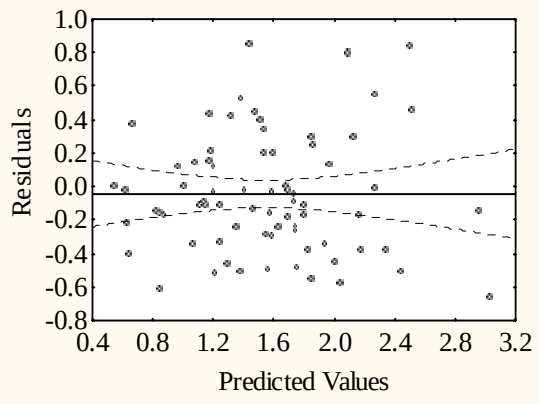

Figure 6.11a: Scatter of residuals. H<170 dataset. Multivariate regression.

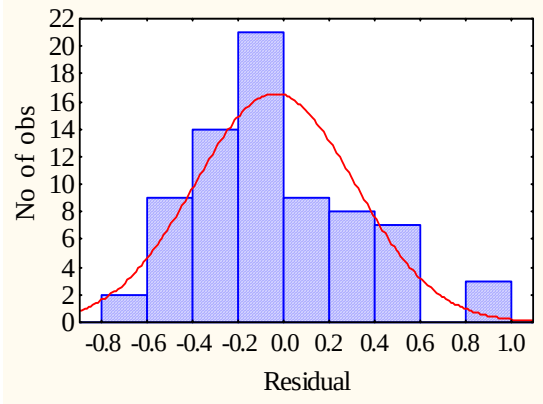

Figure 6.11b: Distribution of residuals. H<170 dataset. Multivariate regression.

6.3.7 Rate of shear strength change

The correlations with the rate of strength change were found to be weak or insignificant (Table 6.5 and 6.6). Consequently, as a first approach, the average loading rates for the snow climates (Table 6.14) were used to calculate the shear strength change over time. The justification for this is:

1. using a large dataset mostly from the intermountain snow climate, Jamieson et al. (2001) found that strength-load ratios averaged 0.98; densification and pressure sintering provide physical explanations for an increase in shear strength with an increased load
2. based on daily average snowfall of 7.8 cm per day from December through March during the winters of 1966 to 1986 at Mt. Fidelity (Schleiss, 1989), and a typical new snow density of 80 kg m^{-3}, the daily average loading rate was 62 Pa d^{-1}. This illustrates the strong effect of load on the shear strength of facet layers in the intermountain snow climate because facet layers at the study sites at Mt. Fidelity and Mt. St. Anne - which

has a snow climate similar to Mt. Fidelity - exhibit an average daily strength increase rate of 58 Pa d^{-1}

3. strength change rates in the Continental data averaged 23 Pa d^{-1} and loading rates at Bow Summit, which was chosen to represent the continental snowpack sites averaged 18 Pa d^{-1} (Johnson, 2000: 60).

It is recognized that while this simple estimate of the strength change rate may include the influence of lagged load (Chalmers, 2001: 79-84), it ignores effects such as temperature, temperature gradient, microstructure and unusually high or low recent loading.

Table 6.14. Loading rates.

Snow climate	Average loading rate (Pa d^{-1})
Continental	18
Intermountain	62
Combined	40

In a second approach, daily loading rates measured at the automatic weather stations in the morning were used. This assumes a lagged load effect, because the precipitation rates were totalled over the period after the last measurement day.

In the next section both the average load and the daily loading rate will be tested in the Forecasting Model.

6.3.8 Model selection

The regression results were not surprising because Load was found to be an important predictor of shear strength in earlier studies (Johnson, 2001: 57-58) and because H and SlabDens are correlated to Load. In the following analysis, Equations 6.3a and 6.3b from the simple regression analysis will be used because:

1. Load was the most significant predictor of shear strength for both continental and intermountain snow climates in Johnson's (2000: 57-58) analysis

2. the effect of Load on shear strength can be explained based on densification (Kojima, 1967; Conway and Wilbour, 1999) and increased pressure sintering
3. the predictors *H* and SlabDens in Table 6.13 were strongly correlated with Load (Johnson, 2000: 59). *H* and SlabDens are not easier to measure than Load and the intercept was not significant (p = 0.11)
4. fitting a multivariate regression to 80 data points for the combined dataset and 61 data points for the H<170 dataset might be questionable
5. the residuals in the multivariate regressions were heteroscedastic
6. logarithmic transformations for multivariate regressions resulted in insignificant B values for Slope and the residuals could not be stabilized for the combined dataset and without the transformation the assumption of homoscedasticity was violated in the combined dataset.

Using the average loading rates for both snow climates (Table 6.14) the Forecasting Model (Equation 6.1) for the combined dataset is re-written as:

Intermountain: $\Sigma_j^* = 1.53\ Load_i^{0.64} + 0.062\ \Delta t_{ij}$ (Equation 6.4a)

Continental: $\Sigma_j^* = 1.53\ Load_i^{0.64} + 0.018\ \Delta t_{ij}$ (Equation 6.4b)

And for the H<170 dataset as:

Intermountain: $\Sigma_j^* = 1.45\ Load_i^{0.53} + 0.062\ \Delta t_{ij}$ (Equation 6.5a)

Continental: $\Sigma_j^* = 1.45\ Load_i^{0.53} + 0.018\ \Delta t_{ij}$ (Equation 6.5b)

Using the daily loading rates (Pcp_{ij}), Equations 6.4 and 6.5 may be expressed:

$$\Sigma_j^* = 1.53\ Load_i^{0.64} + Pcp_{ij} \quad \text{(Equation 6.6a)}$$

and

$$\Sigma_j^* = 1.45\ Load_i^{0.53} + Pcp_{ij} \quad \text{(Equation 6.6b)}$$

respectively.

Table 6.15 summarizes the coefficient of determination and absolute errors between estimated and measured shear strength using Equations 6.3a and 6.3b referred to as the start of the interval and the end of the intervals using Equations 6.4a, 6.4b and 6.5a, 6.5b in the model building dataset for the average loading rates and in Table 6.16 for the daily loading rates.

Table 6.15. Fit of Forecasting Model to data used to build the model using the Power Law model and average loading rates

	Combined		H<170	
Interval	Start N = 101	End N = 82	Start N = 74	End N = 60
Coefficient of determination r^2	0.83	0.79	0.63	0.61
Absolute error (kPa)	0.61	0.67	0.35	0.43

At the start of the interval, the average error for H<170 was about half the value for the combined dataset, which was expected, because the higher variations associated with deeper weak layers were excluded by considering only shallower weak layers. But even though the errors are smaller in the H<170 compared to the combined dataset at the start of the interval the r^2 is significantly worse in the smaller dataset. The same is observed for the end of the intervals.

Table 6.16. Fit of Forecasting Model to data used to build the model using the Power Law model with daily loading rates

	Combined		H<170	
Interval	Start N = 101	End N = 70	Start N = 74	End N = 49
Coefficient of determination r^2	0.84	0.77	0.63	0.58
Absolute error (kPa)	0.61	0.77	0.35	0.49

For the fit of the models using the daily loading rates the same pattern as for the average loading rates is observed. Again the combined dataset seems to have larger errors compared to H<170, although the coefficient of determination is better at both the start and the end of the intervals for the combined dataset.

Zeidler and Jamieson (2002) suggested that the use of daily loading rates in place of average loading rates would most likely improve the ability of the Forecasting Model to predict the shear strength of faceted layers, but comparing the fits at the end of the intervals (Table 6.15 and 6.16) the results do not support this suggestion. Indeed, r^2 for the combined dataset with daily loading rates was 0.77 (compared to 0.79 with average loading rates) and 0.58 (compared to 0.61) for H<170. Also the standard error of the estimates and the coefficients of variation both increased using the daily loading rates. This is possibly because daily loading rates were not available for all the time series and thus the number of observations in the datasets was significantly reduced. However, the lagged response of shear strength to load (Chalmers, 2001: 79-84) and the length of the measurement intervals (mostly 4-8 days) reduce the sensitivity of the shear strength model (but not stability!) to one or two day periods with above or below average loading.

All in all, the models using the combined dataset (Equations 6.4a, b and 6.6a) were the most promising for both deeper and shallower burials. However, a comparison between the combined model and the H<170 model using only data points with H<170 cm was made. Surprisingly, the H<170 model and the combined model predicted the data points where the burial depth did not exceed 170 cm with about the same accuracy. Consequently an independent dataset from an intermountain snowpack will be tested in the next section using only the equations developed with the Combined dataset (Equations 6.4a and 6.6a).

6.3.9 Model testing

One time series from an intermountain snow climate was excluded from the analysis for testing purposes, the facet layer formed at Mt. St. Anne on the 7th of January 2002 (Appendix 1). This time series comprises 12 measurement days over a period of 78 days. In Figure 6.12 the estimated shear strength at the start of an interval (Equation 6.3a) and the shear strength forecast at the end of each interval (Equation 6.4a and Equation 6.6a) are plotted along with measured values. At the start and end of each interval, the Forecasting Models explained 92% of the variability between the observed and predicted results in the test series for both the average and the daily loading rates.

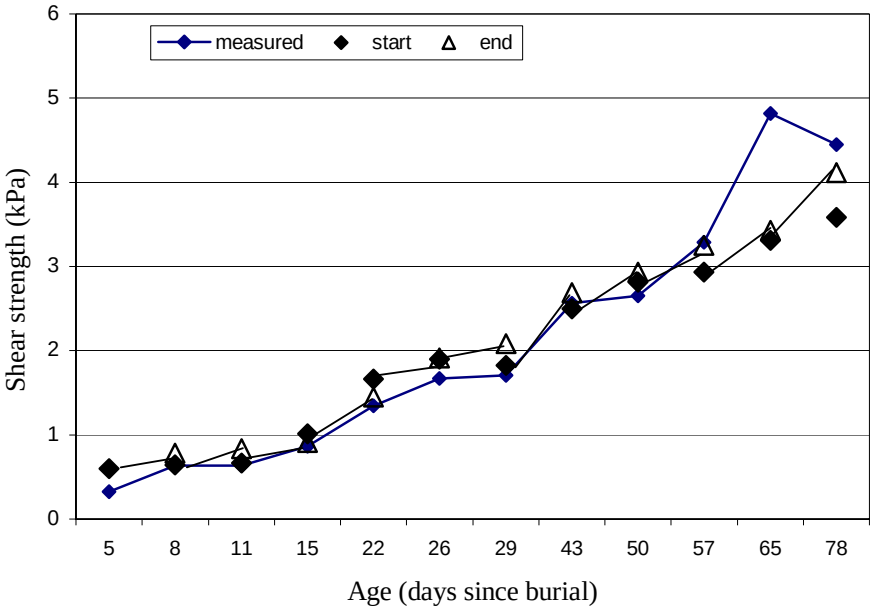

Figure 6.12 a: Model testing: forecasting shear strength at Mt. St. Anne (faceted layer formed 07 January 2002). Average loading rates.

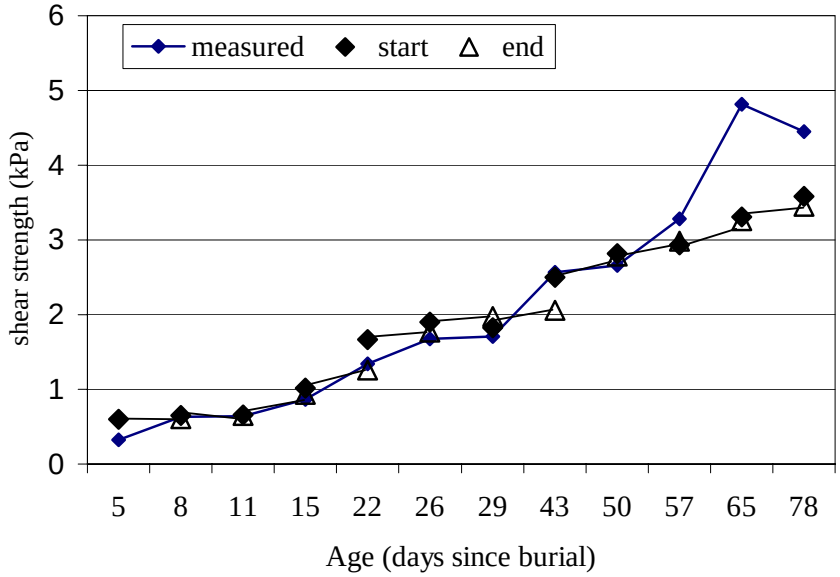

Figure 6.12 b: Model testing: forecasting shear strength at Mt. St. Anne (faceted layer formed 07 January 2002). Daily loading rates.

Up to Day 50, the model using average loading rates overestimated the strength by an average of 9% at the start of the intervals using Equation 6.3a and by 11% (Figure 6.12 a) for the Columbia Mountains (Equation 6.4a). Figure 6.12 b shows the shear strength values at the end of the interval using the daily loading rates. The values fluctuate from an overestimation to underestimation from interval to interval, suggesting that the equation with the daily loading rates is very sensitive. The last three shear strength values in both graphs underestimate the measured strength; however, the effect is not as pronounced when using the average loading rates. At the start of the interval, underestimation is 22% and at the end of the interval, using the average loading rate 14%, and using the daily loading rates 23%. One reason for the high percentage is that on Day 65 (13 March 2002), the models underestimated the measured strength significantly with a maximum of 32% (1.55 kPa) using the daily loading rates, but in view of the better estimate for the next measurement (Day 82), the measured strength on Day 65 may have been higher than average values in the study site due to spatial variability.

The absolute error at the start and end of all forecast intervals was 0.42, 0.46 and 0.47 kPa respectively.

In areas with a snow climate similar to Mt. Fidelity and Mt. St Anne, the model shows potential for estimating shear strength over periods of up to eight days based on a snowprofile (to identify the weak layer and its depth) and load measurement.

Compared to a prior analysis of the dataset, which did not adjust the strength to the normal load (Zeidler and Jamieson, 2002) the prediction improved from 87% to 92% at the start of the interval and from 88% to 92% at the end of the interval.

Both the daily and the average loading rates proved to be valuable for forecasting the shear strength of layers of faceted crystals. However the daily loading rates are preferred in further analysis because the shear strength tends to be overestimated when using the average loading rates. An overestimation implies a prediction of more stable conditions than are observed, which might lead to costly decisions in avalanche terrain. Also using the daily loading rates has the advantage that specific precipitation patterns are reflected in the calculations.

6.4 Surface hoar

6.4.1 Dataset

The dataset for the surface hoar analysis consisted of 76 time series from an intermountain snow climate and 26 from a continental snow climate. Shear frame and snowpack observations were available from the years 1993 to 2004. A time series contained at least three measurement days on one specific weak layer with the crystal form either being reported as 7 (feathery crystals) or 7a (surface hoar crystals) as first or second form (Colbeck at al., 1990). The surface hoar layer buried on the 8th of March 2001 at Mt. St. Anne in the study plot and on an Airbox close to the study plot, a 16 x 3 x 2 m box placed on the ground during the summer. The crystal form on 3 out of 11 measurement days in the study plot and on 3 out of 9 measurement days on the Airbox was reported as 4c in the middle of the series, but the remaining observations again were 7 or 7a as the first form. In total 102 time series with 808 shear strength measurements and 706 shear strength changes fulfilled the requirements, a much bigger dataset than for the facet analysis.

Table 6.17 gives an overview of the time series considered, including the number, the location of the observations and the snow climate. Two series, bolded and italic, were excluded from the model building procedure and are later used for testing purposes, which reduces the dataset to 784 shear strength observations and 684 strength changes.

The predictor and the response variables can be reviewed in Section 6.2.3 and 6.2.4.

Table 6.17: Surface hoar time series. Bolded are time series used for testing purposes. I and C represent intermountain and continental snowpacks, respectively.

WL date	Location	Cli-mate	N	WL date	Location	Cli-mate	N
15-Nov-93	Mt. St. Anne	I	5	07-Jan-95	Rocky Slope	I	7
04-Dec-93	Mt. St. Anne	I	7	07-Jan-95	Vermont	C	11
18-Dec-93	Sam's	I	5	14-Feb-95	Mt. St. Anne	I	5
29-Dec-93	Mt. St. Anne	I	7	28-Dec-95	Bogus	C	14
29-Dec-93	Sam's	I	5	28-Dec-95	MSA cutblock	I	8
05-Feb-94	Mt. St. Anne	I	7	28-Dec-95	Mt. St. Anne	I	11
15-Dec-94	Elk Study Plot	C	9	28-Dec-95	MSA Airbox	I	4
15-Dec-94	Vermont	C	12	28-Dec-95	Pygmy Run	I	8
07-Jan-95	Flaming Corner	I	3	28-Dec-95	Rocky Slope	I	6
07-Jan-95	Mt. St. Anne	I	9	28-Dec-95	Vermont	C	16
07-Jan-95	MSA Airbox	I	8	01-Jan-96	Mt. St. Anne	I	5

Table 6.17: continued

01-Jan-96	MSA Airbox	I	3	30-Dec-99	Fidelity	I	15
04-Feb-96	Mt. St. Anne	I	3	30-Dec-99	Mt. St. Anne	I	13
17-Feb-96	Mt. St. Anne	I	6	31-Jan-00	Cheops	I	9
17-Jan-97	Bogus Study Plot	C	4	31-Jan-00	Fidelity	I	10
17-Jan-97	Mt. St. Anne	I	13	31-Jan-00	Mt. St. Anne	I	12
17-Jan-97	MSA Airbox	I	10	31-Jan-00	MSA Airbox	I	10
17-Jan-97	North Moose Log Cut	C	13	05-Feb-00	Cheops	I	6
17-Jan-97	Pygmy Run	I	8	21-Feb-00	Cheops	I	6
17-Jan-97	Vermont Airbox	C	12	21-Feb-00	Fidelity	I	6
17-Jan-97	Vermont	C	12	17-Nov-00	Fidelity	I	3
10-Feb-97	MSA Cutblock	I	7	22-Nov-00	Mt. St. Anne	I	3
10-Feb-97	***Mt. St. Anne***	I	***9***	24-Nov-00	Fidelity	I	7
10-Feb-97	MSA Airbox	I	8	07-Dec-00	Mt. St. Anne	I	6
11-Feb-97	North Moose Log Cut	C	8	13-Jan-01	Fidelity	I	17
11-Feb-97	Pygmy Run	I	7	20-Jan-01	Mt. St. Anne	I	18
11-Feb-97	Vermont Airbox	C	8	***28-Jan-01***	***Fidelity***	I	***15***
11-Feb-97	Vermont	C	8	23-Feb-01	Fidelity	I	8
08-Dec-97	MSA Cutblock	I	9	23-Feb-01	Mt. St. Anne	I	10
08-Dec-97	Mt. St. Anne	I	7	08-Mar-01	Mt. St. Anne	I	11
26-Dec-97	Middle Moose	C	14	08-Mar-01	MSA Airbox	I	9
26-Dec-97	Vermont	C	7	16-Feb-02	Fidelity	I	7
02-Feb-98	Mt. St. Anne	I	10	16-Feb-02	Mt. St. Anne	I	11
03-Feb-98	Middle Moose	C	13	06-Dec-02	Fidelity	I	3
03-Feb-98	Vermont	C	6	25-Dec-02	Fidelity	I	7
13-Feb-98	Mt. St. Anne	I	9	25-Dec-02	Mt. St. Anne	I	3
13-Feb-98	MSA Airbox	I	11	11-Jan-03	Fidelity	I	6
17-Feb-98	Middle Moose	C	5	19-Jan-03	Mt. St. Anne	I	12
25-Feb-98	Mt. St. Anne	I	7	20-Jan-03	Fidelity	I	8
25-Feb-98	MSA Airbox	I	7	15-Feb-03	Fidelity	I	6
28-Feb-98	Middle Moose	C	5	01-Jan-02	Mt. St. Anne	I	12
28-Feb-98	Vermont Airbox	C	3	02-Jan-02	Fidelity	I	9
28-Feb-98	Vermont	C	3	15-Feb-03	Mt. St. Anne	I	10
03-Jan-99	Cheops	I	4	19-Feb-03	Mt. St. Anne	I	5
03-Jan-99	Fidelity	I	4	16-Nov-03	Fidelity	I	3
24-Jan-99	Cheops	I	3	07-Jan-04	Mt. St. Anne	I	4
24-Jan-99	Fidelity	I	3	14-Feb-04	Fidelity	I	7
16-Feb-99	Cheops	I	10	14-Feb-04	Mt. St. Anne	I	4
16-Feb-99	Fidelity	I	8	24-Feb-04	Fidelity	I	4
30-Dec-99	Bow Summit	C	10	25-Feb-04	Mt. St. Anne	I	5
30-Dec-99	Cheops	I	12	03-Mar-04	Mt. St. Anne	I	4

6.4.2 Descriptive statistics and distribution of the response variables

In this analysis of surface hoar the shear strength as measured with the shear frame was adjusted to normal load, the shear frame size and gravitational forces.

The shear strength (Σ_i) of the 784 measurements ranged from a minimum of 0.07 to a maximum of 8.48 kPa, with a mean of 2.20 kPa. The median indicates that 50% of the measurements are lower than 1.76 kPa. As explained in Section 6.3.2 there is a measurement bias towards more frequent measurements when the weak layer is weaker, which also applies to the observations of buried surface hoar layers. As with the distribution of the shear strength of faceted layers, the shear strength of surface hoar layers is truncated at zero and highly positive skewed (Figure 6.13a). The hypothesis of normality is rejected with the Kolmogorov-Smirnov test ($D = 0.117$, $p < 0.01$) and the Lilliefors test ($p < 0.001$).

The distribution of the shear strength change rate $(\Delta\Sigma/\Delta t)_{ij}$ in Figure 6.13b illustrates that strength loss is less common over time than strength gain, with 125 out of 684 observations (18%) losing strength. The highest strength loss measured was -0.281 kPa per day and the highest strength gain 0.774 kPa per day with a mean of 0.075 kPa per day. The hypothesis of normality is rejected with the Kolmogorv-Smirnov test ($D = 0.085$, $p < 0.01$) and the Lilliefors test ($p < 0.01$).

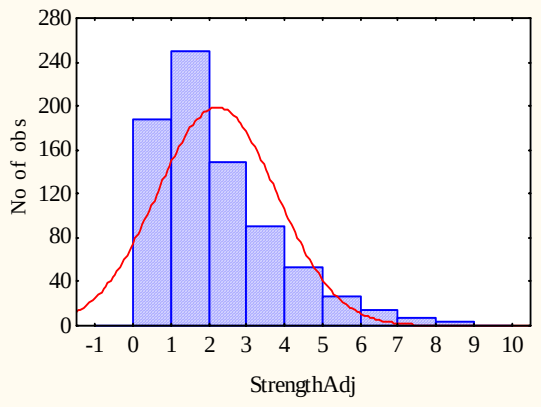

Figure 6.13a: Distribution shear strength.

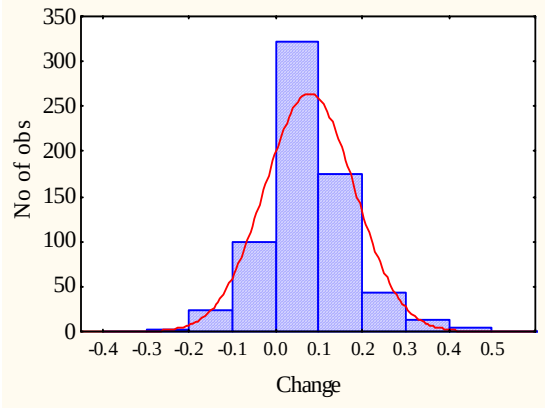

Figure 6.13b: Distribution strength change.

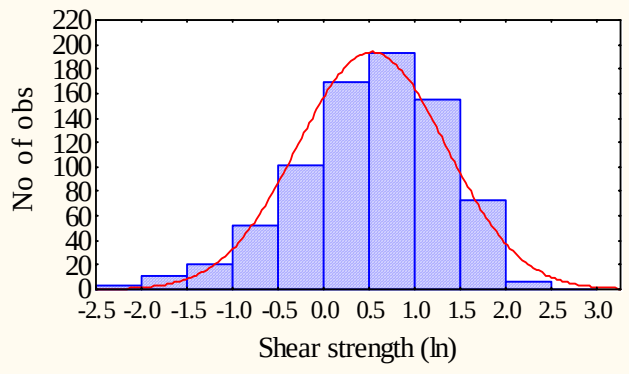

Figure 6.13c: Distribution shear strength (logarithmic).

The normality is not rejected with the Kolmogorov-Smirnov test ($D = 0.044$, $p < 0.1$) but with the Lilliefors test ($p < 0.01$) in the logarithmic transformation of shear strength (Figure 6.13c).

6.4.3 Exploring the dataset

Similar to the analysis of faceted layers the entire dataset was assessed for its suitability to predict skier-triggered avalanches. However, the surface hoar dataset set is larger and a better split into subsets more or less relevant to skier triggering might be possible.

In his interval model, Chalmers (2001) argued that buried surface hoar layers are most likely to cause skier-triggered slab avalanches within the first 30 days after burial, and therefore excluded measurements older than 30 days. However, Figure 6.14a shows that a cut-off at an age of 30 days ignores many observations with low shear strength values (< 2.5 kPa), possibly ignoring important information for years with below average snowpack height years. Another approach involves considering the depth of the weak layers, as proposed in Section 6.3.3 for layers of faceted crystals. As mentioned in Section 1.3, skier triggering is more likely at depth less than 100 cm (Föhn and Camponovo, 1997), valuable information that could help to improve the forecast. In Figure 6.14b the shear strength is plotted versus the slab thickness. Comparing both figures the scatter of Figure 6.14a is greater than in Figure 6.14b suggesting that slab thickness is a better predictor than age. Also the second graph shows that at a depth greater than 80-100 cm the variance of

shear strength increases. A cut-off at 100 cm was chosen, instead of 80 cm, because it might better capture the transition between unstable and stable conditions. In addition, the dataset consists of series from an intermountain snow climate as well as from a continental snowpack. The shear strength and shear strength change rate of a surface hoar layer might be strongly influenced by the different snow climates, for example the expected higher temperature gradients in the continental snowpack or the higher loads in the intermountain snowpack. Considering these differences, four different datasets will be analyzed:

1) the entire dataset (Entire),
2) dataset with a maximum slab thickness of 100 cm, intermountain and continental ($H<100_{comb}$),
3) dataset with a maximum slab thickness of 100 cm, only intermountain ($H<100_{inter}$), and
4) dataset with a maximum slab thickness of 100 cm, only continental ($H<100_{cont}$).

This allows for a comparison of the datasets, and to determine whether the intermountain snow climate and the continental snow climate have different shear strength behaviors, and if deeper and shallower layers can be calculated with one model.

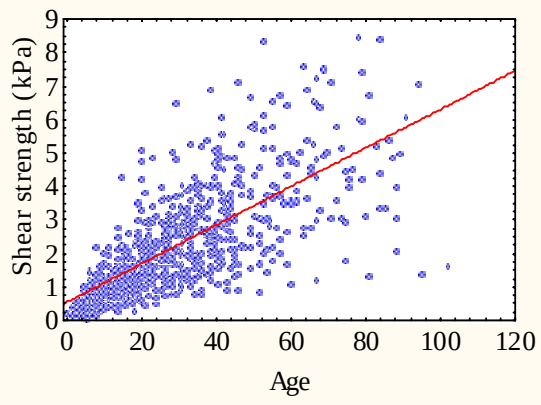

Figure 6.14a: Scatterplot of number of days since burial (Age) vs shear strength

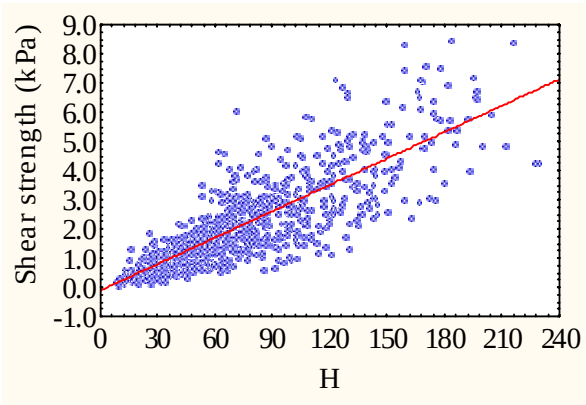

Figure 6.14b: Scatterplot of slab thickness (H) vs shear strength

A partition of the dataset that uses load as predictor could be potentially useful, but was not made because load correlates highly with slab thickness and the latter is easier to measure. Table 6.18 gives an overview of the descriptive statistics for shear strength of the four different datasets.

Table 6.18. Descriptive statistics of strength for the four datasets in the surface hoar analysis.

Dataset	Entire	H<100$_{comb}$	H<100$_{inter}$	H<100$_{cont}$
N	784	581	394	187
Minimum (kPa)	0.065	0.065	0.104	0.065
Maximum (kPa)	8.483	6.084	6.084	4.601
Mean (kPa)	2.198	1.576	1.654	1.411
Median (kPa)	1.760	1.380	1.510	1.218
Variance (kPa^2)	2.477	0.964	1.007	0.840
Standard Deviation (kPa)	1.574	0.982	1.003	0.916
Standard Error (kPa)	0.056	0.041	0.051	0.067

As expected, the entire dataset has the highest variance, because deeper burials are included. In the continental snowpack the lowest minimum shear strength was measured and the lowest maximum shear strength compared to the other datasets. However, the continental dataset consists of less than half the number of observations as the intermountain dataset.

6.4.4 Correlation analysis

Shear strength and shear strength change rate as shown in Section 6.4.2 are not normally distributed, and consequently Spearman rank correlation analysis was applied to assess the importance of the predictors. Table 6.19 shows the results for the entire dataset, Table 6.20 for the H<100$_{comb}$ dataset, Table 6.21 for the H<100$_{inter}$ dataset and Table 6.22 for the H<100$_{cont}$ dataset correlating the predictors (Table 6.1) and shear strength.

Table 6.19. Spearman rank correlations with shear strength for the Entire dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
Load	781	0.917	<1E-17
H	784	0.836	<1E-17
SlabDens	781	0.790	<1E-17
Age	784	0.782	<1E-17
HS	782	0.456	<1E-17
T –5	782	0.427	<1E-17
Thick	656	-0.240	4.7E-10
Emax	778	-0.206	7.0E98
Emin	778	-0.136	1.4E-4
TG	782	0.115	0.001
Slope	*784*	*0.029*	*0.414*

Table 6.20. Spearman rank correlations with shear strength for the $H<100_{comb}$ dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
Load	580	0.887	<1E-17
SlabDens	580	0.763	<1E-17
H	582	0.747	<1E-17
Age	582	0.738	<1E-17
Thick	496	-0.438	<1E-17
T –5	581	0.394	<1E-17
Emax	577	-0.350	<1E-17
Emin	577	-0.274	2.3E-11
HS	582	0.214	1.9E-7
Slope	*582*	*-0.068*	*0.100*
TG	*581*	*0.035*	*0.398*

Table 6.21. Spearman rank correlations with shear strength for the H<100$_{inter}$ dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
Load	392	0.881	<1E-17
Age	394	0.830	<1E-17
SlabDens	392	0.794	<1E-17
H	394	0.702	<1E-17
T –5	393	0.484	<1E-17
Thick	343	-0.465	<1E-17
Emax	390	-0.377	1.3E-14
Emin	390	-0.335	1.2E-11
HS	394	0.177	4.3E-4
Slope	394	-0.120	0.017
TG	*393*	*0.022*	*0.672*

Table 6.22. Spearman rank correlations with shear strength for the H<100$_{cont}$ dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
Load	187	0.885	<1E-17
H	187	0.792	<1E-17
Age	187	0.787	<1E-17
SlabDens	187	0.758	<1E-17
Thick	152	-0.460	2.4E-9
Emax	186	-0.391	3.3E-8
Emin	186	-0.380	8.9E-8
T –5	187	0.267	2.2E-4
HS	187	0.207	0.005
Slope	*187*	*0.060*	*0.415*
TG	*187*	*-0.025*	*0.730*

The four highest ranked predictor variables in all of the Spearman rank correlations with shear strength were load, slab density, slab thickness, and age, similar to the highly ranked variables in the facet analysis. The explanations for these predictor variables were summarized in Section 6.3.4, except age, which is directly related to Load, H and SlabDens

and suggests that the older a surface hoar layer is, the stronger it tends to be, which is not surprising. However, the physical effect of Age is likely weaker than Load and related variables, especially in non-average snowpack years where the snowpack height might be unexceptional high or low.

The temperature gradient over the weak layer was only significant in the entire dataset. This suggests that for deeper burials the temperature gradient is more important, possibly because a higher temperature gradient preserves the crystals for a longer time. Similarly, Chalmers (2001) found that load, slab thickness, and age mostly correlated with shear strength, though he did not consider slab density in his analysis.

Table 6.23 shows the results for the entire dataset, Table 6.24 for the $H<100_{comb}$ dataset, Table 6.25 for the $H<100_{inter}$ dataset and Table 6.26 for the $H<100_{cont}$ dataset correlating the predictors (Table 6.1) and shear strength change rate.

Table 6.23. Spearman rank correlations with shear strength change rate for the entire dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
TG	682	0.254	1.8E-11
SlabDens	681	-0.212	2.4E-8
Age	684	-0.204	7.7E-8
HS	682	0.180	2.3E-6
Strength	684	-0.139	2.8E-4
T -5	*682*	*-0.051*	*0.183*
H	*684*	*0.050*	*0.192*
Load	*681*	*-0.041*	*0.291*
Thick	*573*	*-0.038*	*0.367*
Slope	*684*	*0.034*	*0.379*
Emax	*678*	*-0.027*	*0.476*
Emin	*678*	*-0.009*	*0.821*

Table 6.24. Spearman rank correlations with shear strength change rate for the $H<100_{comb}$ dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
TG	532	0.306	5.0E-13
Age	533	-0.238	2.7E-8
HS	533	0.234	4.4E-8
SlabDens	531	-0.230	8.3E-8
Strength	533	-0.162	1.7E-4
H	533	0.095	0.029
Slope	*533*	*0.062*	*0.153*
Load	*531*	*-0.035*	*0.420*
T –5	*532*	*-0.028*	*0.525*
Emax	*528*	*-0.024*	*0.580*
Emin	*528*	*0.018*	*0.684*
Thick	*451*	*-0.018*	*0.709*

Table 6.25. Spearman rank correlations with shear strength change for the $H<100_{inter}$ dataset. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
TG	364	0.363	9.2E-13
SlabDens	363	-0.277	8.2E-8
Age	365	-0.270	1.6E-7
Strength	365	-0.224	1.5E-5
HS	365	0.162	0.002
Load	363	-0.107	0.043
Slope	*365*	*0.099*	*0.060*
T -5	*364*	*-0.054*	*0.301*
Emin	*361*	*-0.047*	*0.376*
H	*365*	*0.036*	*0.493*
Emax	*361*	*-0.024*	*0.644*
Thick	*316*	*-0.021*	*0.715*

Table 6.26. Spearman rank correlations with shear strength change for the $H<100_{cont}$ dataset. Insignificant variables ($p > 0.05$) are in italics.

Variable	Valid N	Spearman R	p-level
Strength	168	-0.154	0.046
TG	*168*	*0.125*	*0.106*
SlabDens	*168*	*-0.124*	*0.111*
Emax	*167*	*-0.098*	*0.209*
H	*168*	*0.080*	*0.300*
Age	*168*	*-0.076*	*0.329*
T –5	*168*	*0.055*	*0.478*
HS	*168*	*0.037*	*0.639*
Thick	*135*	*-0.336*	*0.737*
Emin	*167*	*0.011*	*0.884*
Load	*168*	*0.009*	*0.907*
Slope	*168*	*-0.007*	*0.928*

The Spearman rank correlations with shear strength change rate revealed different results for the continental dataset, where only the shear strength of the weak layer on the last measurement day showed a weak negative correlation. In at least two out of the other three datasets, five of the fourteen predictor variables were significantly correlated with the rate of shear strength change:

- **TG:** The highest ranked predictor for the strength change rate in all datasets was the temperature gradient measured over 10 cm across the failure plane of the weak layer. Since most measurements had a negative TG, the positive correlation suggests that the closer the temperature gradient was to 0°C/m the greater is the change in strength of the surface hoar layer. High magnitude of temperature gradient (farther from 0°C/m) is often associated with cold temperatures where bonding slows down, whereas bonding is promoted when the temperature gradient is small, as often observed when the snowpack and the slab are thicker.
- **HS:** The snowpack height was significantly and positively correlated in all three datasets with the strength change rate. The positive relationship

implies that higher shear strength change rates are associated with higher snow depth.

- **SlabDens:** The slab density was significant in all datasets with a negative correlation, suggesting that the rate in change is greater when the slab is less dense. This is likely because the shear strength of a surface hoar layer changes more quickly when they are closer to the surface and the weak layer has little densified.
- **Age:** The age of the weak layer exhibits a negative correlation in all datasets with the rate of change. This can be related to the slab density where younger layers gain shear strength faster when they are shallower right after their burial.
- **Strength:** The initial shear strength is negatively correlated with the shear strength change rate. This implies that weaker surface hoar layers likely gain shear strength faster.

6.4.5 Simple linear regression

The correlation analysis revealed that load was the highest ranked predictor variable for shear strength and will therefore be used in the simple regression analysis. The procedure is outlined in Section 6.2.2 and was already used in the analysis of faceted layers.

Instead of analyzing the outliers per dataset, an overview will be given of their physical causes, because the outliers for each dataset were similar or at least followed a specific pattern.

Table 6.27 summarizes the simple regression results after removing the outliers for each dataset.

Table 6.27: Simple linear regression of strength on Load.

Dataset	N	r^2	Standard Error of Estimate	CV	p	Kolmogorov-Smirnov D	Kolmogorov-Smirnov p	Lilliefors p
Entire	752	0.85	0.572	28	<1.0E-17	0.063	<0.01	<0.009
$H<100_{comb}$	570	0.77	0.444	29	<1.0E-17	0.063	<0.05	<0.009
$H<100_{inter}$	382	0.75	0.455	29	<1.0E-17	0.068	<0.1	<0.01
$H<100_{cont}$	185	0.79	0.403	29	<1.0E-17	0.080	<0.2	<0.01

The coefficients of determination for all the datasets ranged from 0.75 to 0.85, a quite promising result. The intermountain dataset had the poorest fit and the second highest standard error, though the coefficients of variation are similar. The highest standard error was observed in the Entire dataset, which could be expected because higher shear strength values, which exhibit higher variances, were included. However Figures 6.15a-d reveal non-constant residual variance, which violates the assumption of homoscedasticity of the residuals in regression analysis. In addition, the assumption of the normal distribution of the residuals was violated in all datasets with the Kolmogorov-Smirnov and the Lilliefors test (Table 6.27).

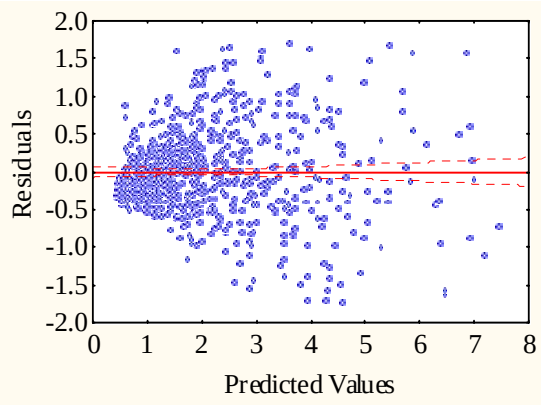

Figure 6.15a: Scatterplot of residuals. Entire dataset.

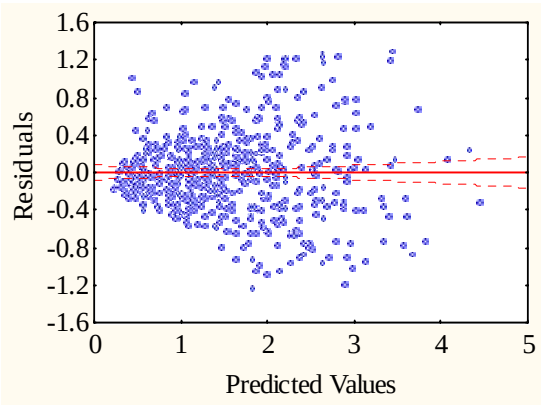

Figure 6.15b: Scatterplot of residuals. $H<100_{comb}$ dataset.

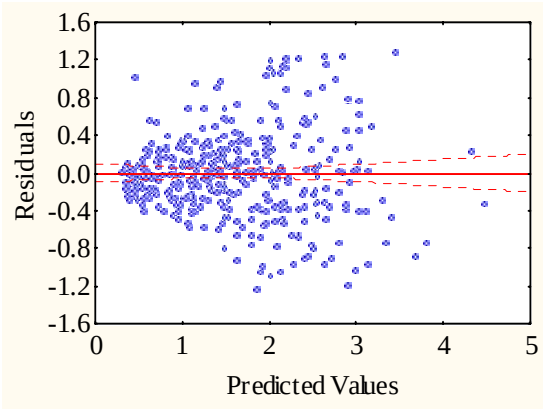

Figure 6.15c: Scatterplot of residuals. $H<100_{inter}$ dataset.

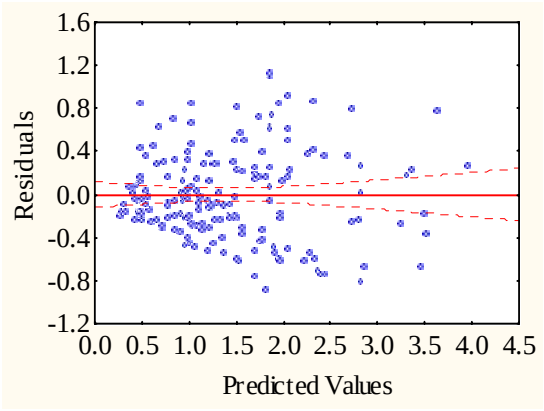

Figure 6.15d: Scatterplot of residuals. $H<100_{cont}$ dataset.

As in the analysis for faceted layers, a logarithmic transformation (Equation 6.2a) was performed because the variance of the residuals in the simple linear regression was not constant. However, for ease of interpretation, the logarithmic equation must be transformed to the power law model (Equation 6.2b).

Table 6.28 lists the regression results of the Entire dataset after the logarithmic transformation and the removal of the outliers.

Table 6.28: Simple logarithmic regression. Entire dataset.

N = 746	**B**	**Standard Error B**	**Significance Level *p***	**r^2**	**Standard Error of Estimate**	***p***
Intercept	0.393	0.011	< 1.0E-17	0.86	0.289	< 1.0E-17
ln Load	0.887	0.013	< 1.0E-17			

The r^2 value improved from 0.85 to 0.86, and the residuals followed a normal distribution using the Kolomogorov-Smirnov test (D = 0.029, p < n.s.) and the Lilliefors test (p < 0.15) (Figure 6.16b). However, even though the distribution of the variance was improved (Figure 6.16a) compared to Figure 6.15a, the scatter of the residuals was higher for lower predicted logarithmic values.

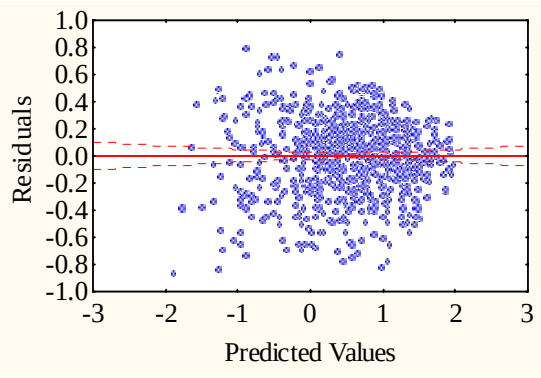

Figure 6.16a: Scatter of residuals from the regression of ln shear strength on ln Load. Entire dataset.

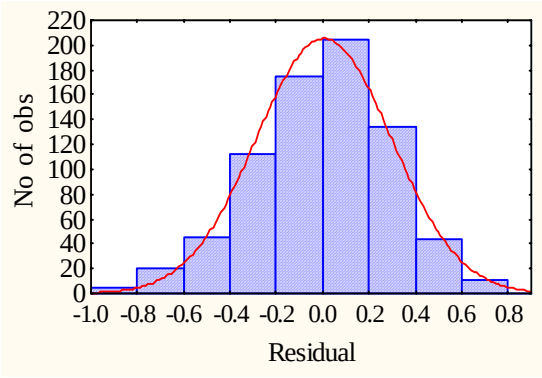

Figure 6.16b: Distribution of residuals from the regression of ln shear strength on ln Load. Entire dataset.

A similar pattern was observed in the $H<100_{comb}$ dataset where the r^2 value improved from 0.77 to 0.82 (Table 6.29), the variance was partly stabilized (Figure 6.17a) and the normal distribution was not rejected with the Kolmogorov-Smirnov test (D = 0.04, p < n.s.) but with the Lilliefors test (p < 0.05) (Figure 6.17b).

Table 6.29: Simple logarithmic regression. H<100 dataset

N = 565	B	Standard Error B	Significance Level *p*	r^2	Standard Error of Estimate	*p*
Intercept	0.412	0.013	< 1.0E-17	0.82	0.3	< 1.0E-17
Load	0.926	0.018	< 1.0E-17			

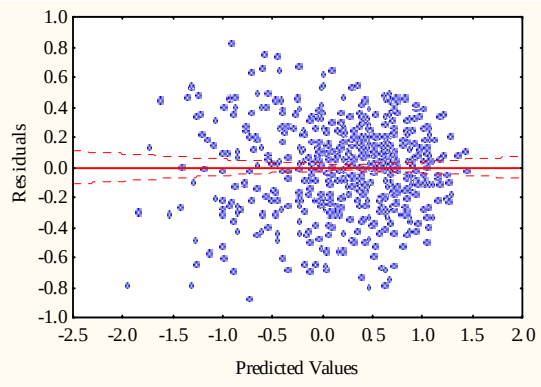

Figure 6.17a: Scatter of residuals from the regression of ln shear strength on ln Load. H<100 dataset.

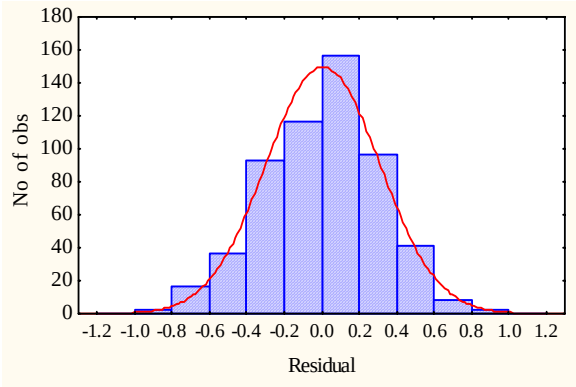

Figure 6.17b: Distribution of residuals from the regression of ln shear strength on ln Load. H<100 dataset.

In addition, the fit of the $H<100_{inter}$ dataset improved from an r^2 value of 0.75 to 0.82 (Table 6.30), the variance was partly stabilized (Figure 6.18a) and the normal distribution was not rejected with the Kolmogorov-Smirnov test (D = 0.06, p<0.15) but with the Lilliefors test (p<0.01) (Figure 6.18b).

Table 6.30: Simple logarithmic regression. $H<100_{inter}$ dataset.

N = 378	**B**	**Standard Error B**	**Significance Level *p***	**r^2**	**Standard Error of Estimate**	***p***
Intercept	0.421	0.016	< 1.0E-17	0.82	0.293	< 1.0E-17
Load	0.934	0.023	< 1.0E-17			

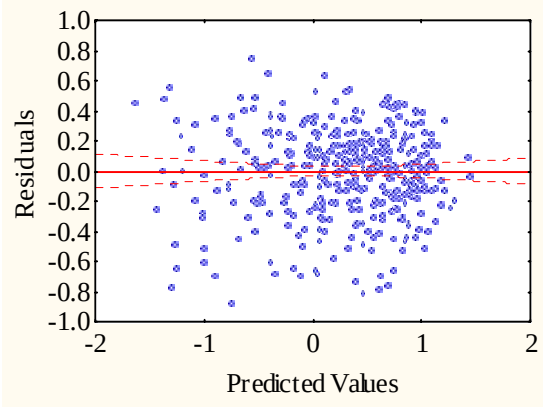

Figure 6.18a: Scatter of residuals from the regression of ln shear strength on ln Load. $H<100_{inter}$ dataset.

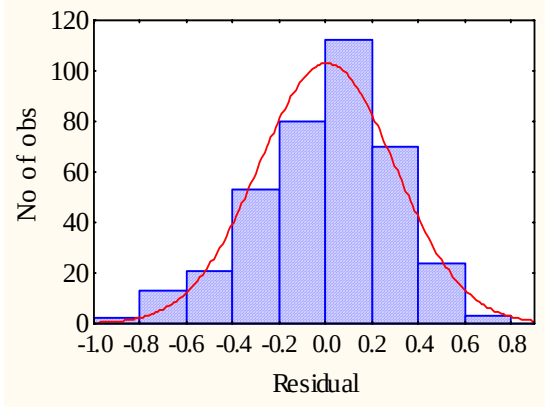

Figure 6.18b: Distribution of residuals from the regression of ln shear strength on ln Load. $H<100_{inter}$ dataset.

The result for the $H<100_{cont}$ dataset also shows an improvement of the r^2 value from 0.79 to 0.82 (Table 6.31) and the variance was partly stabilized (Figure 6.19a). The normal distribution of the residuals was not rejected with the Kolmogorov-Smirnov test (D = 0.065, p < n.s.) and with the Lilliefors test (p < 0.1) (Figure 6.19b).

Table 6.31: Simple logarithmic regression. H<100$_{cont}$ dataset						
N = 186	**B**	**Standard Error B**	**Significance Level *p***	**r^2**	**Standard Error of Estimate**	***p***
Intercept	0.402	0.025	< 1.0E-17	0.82	0.316	< 1.0E-17
Load	0.910	0.032	< 1.0E-17			

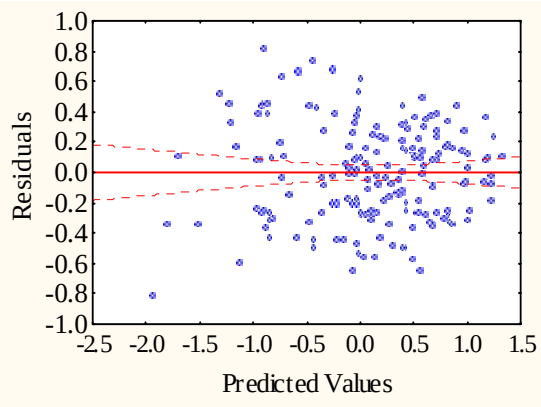

Figure 6.19a: Scatter of residuals from the regression of ln shear strength on ln Load. H<100$_{cont}$ dataset.

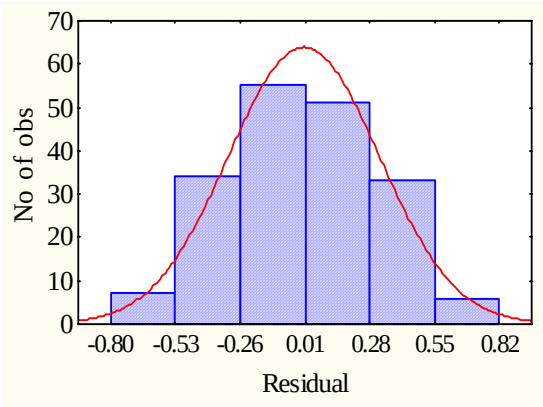

Figure 6.19b: Distribution of residuals from the regression of ln shear strength on ln Load. H<100$_{cont}$ dataset.

As mentioned above, the logarithmic transformations have to be back-transformed in order to give a meaningful prediction for shear strength. The logarithmic equations may be re-written as a power law (see Equation 6.2b):

Entire dataset:

$$\Sigma = 1.48\ Load^{0.89} \qquad \textit{Equation 6.7a}$$

H<100$_{comb}$ dataset:

$$\Sigma = 1.51\ Load^{0.93} \qquad \textit{Equation 6.7b}$$

H<100$_{inter}$ dataset:

$$\Sigma = 1.52\ Load^{0.93} \qquad \textit{Equation 6.7c}$$

H<100$_{cont}$ dataset:

$$\Sigma = 1.50\ Load^{0.91} \qquad \textit{Equation 6.7d}$$

In all of the above datasets the coefficients are similar and for the simple regression analysis a separation of the datasets seems to offer no advantage for the prediction of shear strength of surface hoar layers.

Altogether 50 different days were identified as outliers in at least one out of the four datasets, though most outliers occurred in more than one dataset. Most of the outliers yielded unusually high shear strength values compared to the shear strength normally measured at the same load. The analysis of the outliers is in close conformity with Chalmers' (2001: 69) findings. For example, measurements during the winter of 1998/1999 at Cheops and Mt. Fidelity showed higher shear strength values associated with a comparable load in other years, though a real physical cause could not be identified and a constant error in the measurements is expected. Often measurements at the end of March and beginning of April were statistical outliers with either very high shear strength values reported or sometimes very low shear strength values. Warmer temperatures may have provoked rapid strengthening of the surface hoar layers, whereas the low values might be associated with spatial variability, e.g. pockets of well preserved crystals. Not surprisingly, older layers were often outliers, which was expected because of the increasing variance. However, through the logarithmic transformation some tests in very young layers (1-3 days) became outliers, likely because a logarithmic transformation reduces the lower values less than higher values (Wilks 1995: 172).

Though the results from the simple regression analysis with load as a predictor seem to be quite promising, multivariate regression analysis was performed to find a "better" model, which might involve physical interactions between predictor variables.

6.4.6 Multiple linear regression

As in Section 6.3.6 in Tables 6.32a to d the strongly correlated variables ($r > 0.8$) are bolded. Again Emax is excluded from further analysis though it is not above 0.8, but Emax was recorded inconsistently. H and Load are the highest correlated especially in the continental and the entire dataset.

Table 6.32a: Correlation between predictors (r). Entire dataset. Surface hoar. Bolded are strong correlation (r > 0.8).											
Variables	**Age**	**Slope**	**E min**	**E max**	**Thick**	**T-5**	**HS**	**H**	**Load**	**TG**	**Slab Dens**
Age	1										
Slope	0.04	1									
Emin	-0.02	0.06	1								
Emax	-0.02	0.16	**0.81**	1							
Thick	-0.01	0.07	0.61	0.72	1						
T-5	0.48	0.07	-0.03	-0.06	-0.16	1					
HS	0.16	0.25	0.04	0.03	-0.03	-0.13	1				
H	0.62	0.19	-0.01	-0.03	-0.02	0.31	0.60	1			
Load	0.74	0.08	-0.04	-0.06	-0.04	0.36	0.51	**0.94**	1		
TG	0.05	0.05	0.02	0	-0.04	0.25	0.24	0.15	0.13	1	
Slab Dens	0.75	-0.16	-0.12	-0.18	-0.17	0.42	0.18	0.55	0.74	0	1

Table 6.32b: Correlation between predictors (r). H<100$_{comb}$ dataset. Surface hoar. Bolded are strong correlation (r > 0.8).											
Variables	**Age**	**Slope**	**E min**	**E max**	**Thick**	**T-5**	**HS**	**H**	**Load**	**TG**	**Slab Dens**
Age	1										
Slope	-0.03	1									
Emin	-0.07	0.05	1								
Emax	-0.08	0.12	**0.82**	1							
Thick	-0.20	0.09	0.62	0.73	1						
T-5	0.45	0.04	-0.06	-0.07	-0.16	1					
HS	-0.13	0.22	0.04	0	-0.08	-0.27	1				
H	0.47	0.15	-0.09	-0.15	-0.15	0.28	0.37	1			
Load	0.68	-0.02	-0.15	-0.22	-0.21	0.37	0.23	**0.87**	1		
TG	-0.03	0.04	0.01	0	-0.03	0.24	0.21	0.11	0.08	1	
Slab Dens	0.70	-0.22	-0.18	-0.25	-0.26	0.37	-0.04	0.42	0.78	-0.07	1

Table 6.32c: Correlation between predictors (r). H<100$_{inter}$ dataset. Surface hoar. Bolded are strong correlation (r > 0.8).

Variables	Age	Slope	E min	E max	Thick	T-5	HS	H	Load	TG	Slab Dens
Age	1										
Slope	-0.07	1									
Emin	-0.21	0.06	1								
Emax	-0.22	0.17	**0.84**	1							
Thick	-0.29	0.13	0.63	0.72	1						
T-5	0.52	0.15	-0.13	-0.12	-0.18	1					
HS	-0.02	0.18	-0.05	-0.06	-0.17	-0.26	1				
H	0.56	0.16	-0.14	-0.16	-0.17	0.39	0.30	1			
Load	0.77	-0.06	-0.21	-0.26	-0.23	0.46	0.19	**0.85**	1		
TG	0.02	0.03	-0.05	-0.06	-0.08	0.17	0.19	0.09	0.06	1	
Slab Dens	0.72	-0.31	-0.26	-0.32	-0.27	0.39	-0.02	0.39	0.79	-0.06	1

Table 6.32d: Correlation between predictors (r). H<100$_{cont}$ dataset. Surface hoar. Bolded are strong correlation (r > 0.8).

Variab	Age	Slope	E min	E max	Thick	T-5	HS	H	Load	TG	Slab Dens
Age	1										
Slope	0.01	1									
Emin	0.30	0.01	1								
Emax	0.17	0.01	0.77	1							
Thick	0.02	-0.04	0.56	0.77	1						
T-5	0.39	-0.25	0.16	0.07	-0.12	1					
HS	0.13	0.59	-0.05	-0.02	0.02	-0.33	1				
H	0.63	0.12	-0.11	-0.21	-0.23	0.11	0.39	1			
Load	0.78	0.09	-0.06	-0.18	-0.24	0.23	0.26	**0.92**	1		
TG	0.04	0.07	0.03	0.06	0.03	0.40	0.05	0.04	0.05	1	
Slab Dens	0.78	0.01	0.06	-0.08	-0.20	0.33	0.02	0.55	0.79	-0.06	1

Stepwise regression analysis was chosen because some predictor variables may be correlated and consequently the full set of potential predictors contains redundant information (Wilks, 1995: 188). In addition, backward elimination of the variables was chosen because in the forward elimination some B values of the predictors and intercepts were insignificant. Again, observations were removed when the standard residuals had a value of less than –3.0 or higher 3.0, as shown in Figure 6.20. The valid number of observations in each dataset depended on the availability of all predictors. Unfortunately, not all variables, for example the thickness of the weak layer, were measured for all values of shear strength.

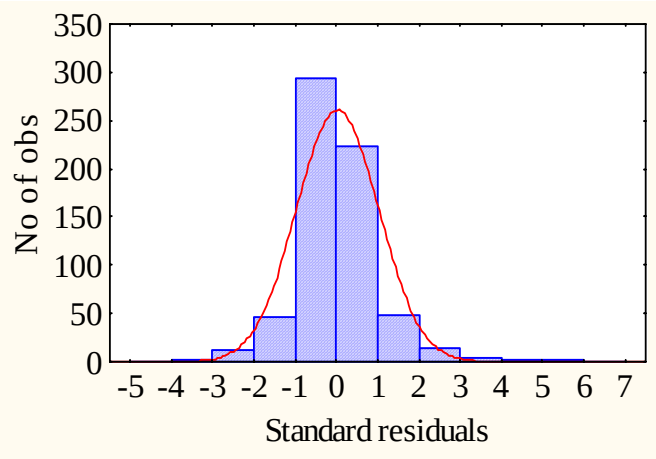

Figure 6.20: Normality plot of standard residuals. Values less than –3.0 or greater 3.0 are statistical outliers.

The results from the multiple stepwise regression using backwards elimination of the predictor variables for the Entire dataset are summarized in Table 6.33. Thick, T–5, HS, Load and SlabDens were selected, after 27 outliers were revealed and deleted. The B value for the intercept was not significant. The Kolmogorov-Smirnov test ($D = 0.069$, $p < 0.01$) and the Lilliefors test ($p < 0.009$) indicate that the residuals are not normally distributed (Figure 6.21b). In addition to the violation of the assumption of the normal distribution the residuals are also heteroscedastic (Figure 6.21a).

Table 6.33: Multiple stepwise (backwards) regression of shear strength. Entire dataset.								
N = 620	**B**	**Standard Error B**	**Signif. Level *p***	**Tolerance**	**Adj. r^2**	**Standard Error of Estimate**	**CV %**	***p***
Intercept	0.095	0.128	0.457		0.88	0.498	24	<1.0E-17
Thick	-0.297	0.036	4.1E-16	0.93				
T-5	0.055	0.013	3.7E-5	0.74				
HS	0.001	0.0003	9.2E-6	0.60				
Load	0.964	0.034	<1.0E-17	0.30				
SlabDens	0.004	0.0006	4.2E-11	0.41				

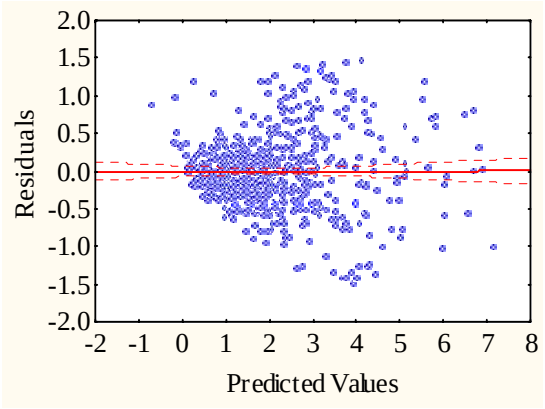

Figure 6.21a: Scatterplot of residuals. Multiple regression. Entire dataset.

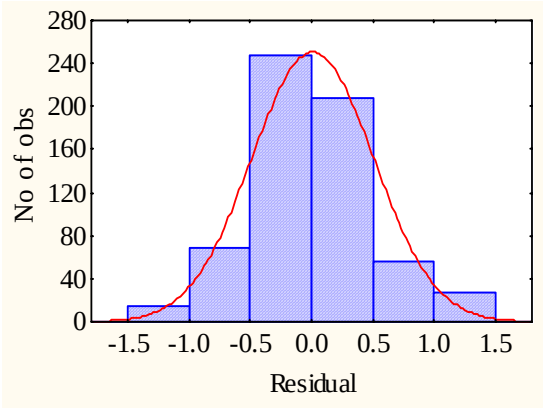

Figure 6.21b: Distribution of residuals. Multiple regression. Entire dataset.

Table 6.34 shows the results of the multiple stepwise regression for the $H<100_{comb}$ dataset after 17 outliers were deleted. The selected predictor variables are similar to the ones picked in the Entire dataset, except that the slab density (SlabDens) is replaced by the slab thickness (H). The r^2 value indicates a less good fit than in the Entire dataset (0.83 compared to 0.88), however all B values are significant. Although the variance of the residuals does not increase as dominantly as in the Entire dataset, a trend of increasing variance with an increase of the predicted value is still visible (Figure 6.22). However, the

normality of the residuals is not rejected with the Kolmogorov-Smirnov test (D = 0.041, p < n.s.) and the Lilliefors test (p < 0.1) (Figure 6.22b).

Table 6.34: Multiple stepwise (backwards) regression of strength. H<100 dataset

N = 473	B	Standard Error B	Signifi. Level *p*	Tolerance	Adj. r^2	Standard Error of Estimate	CV %	*p*
Intercept	0.642	0.071	<1.0E-17		0.83	0.364	24	<1.0E-17
Thick	-0.265	0.029	<1.0E-17	0.94				
T –5	0.044	0.011	3.4E-5	0.77				
HS	0.001	0.0002	1.1E-5	0.73				
H	-0.009	0.002	2.3E-8	0.21				
Load	1.573	0.065	<1.0E-17	0.21				

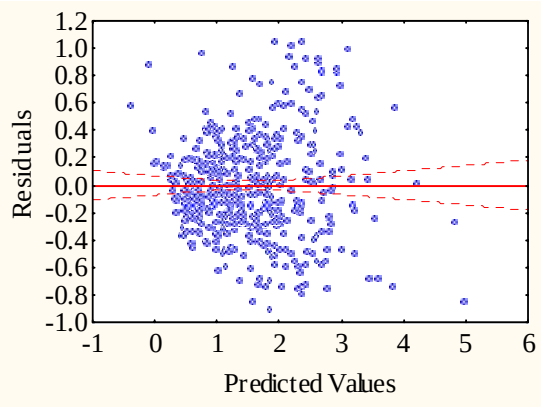

Figure 6.22a: Scatterplot of residuals. Multiple regression. H<100 dataset.

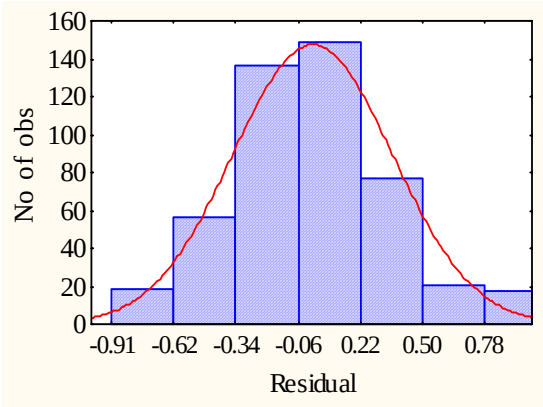

Figure 6.22b: Distribution of residuals. Multiple regression. H<100 dataset.

The $H<100_{inter}$ dataset, with an r^2 of 0.84 (Table 6.35) had a less good fit than the Entire dataset, though the same variables were selected by the backward stepwise regression. But also the B value of the intercept was not significant (p = 0.47). Fourteen observations were deleted because they were outliers. Again the scatter of the residuals increased with increasing predicted values (Figure 6.23a). The normality of the residuals

was not rejected with the Kolmogorov-Smirnov test (D = 0.054, p < n.s.), but was rejected with the Lilliefors test (p < 0.05).

Table 6.35: Multiple stepwise (backwards) regression. H<100$_{inter}$ dataset

N = 325	B	Standard Error B	Signif. Level *p*	Tolerance	Adj. r^2	Standard Error of Estimate	CV %	*p*
Intercept	0.091	0.125	0.47		0.84	0.367	23	<1.0E-17
Thick	-0.216	0.033	2.1E-10	0.88				
T-5	0.063	0.014	6.8E-6	0.68				
HS	0.001	0.0003	4.5E-5	0.74				
Load	0.906	0.068	<1.0E-17	0.29				
SlabDens	0.004	0.0006	4.9E-10	0.34				

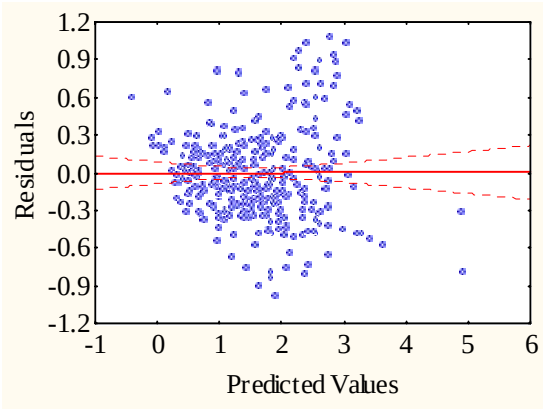

Figure 6.23a: Scatterplot of residuals. Multiple regression. H<100$_{inter}$ *dataset.*

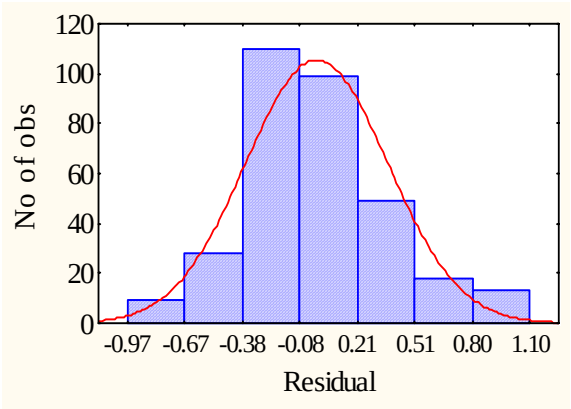

Figure 6.23b: Distribution of residuals. Multiple regression. H<100$_{inter}$ *dataset.*

In the H<100$_{cont}$ dataset the predictor variables Thick, H and Load were selected by the backward elimination in the stepwise regression with only one outlier removed. The results are summarized in Table 6.36. As already observed in the other datasets, Load is the most important predictor and yields the highest and most influential B value. The scatter of the residuals is heterorscedastic (Figure 6.23a). The Lilliefors test (p<0.1) and the Kolmogorov-Smirnov test (D = 0.05, p<n.s.) do not reject the hypothesis that the residuals

are normally distributed (Figure 6.23b). The tolerance in Table 6.36 is lower than 0.2 for H and Load indicating a serious problem with multicollinearity. The continental dataset is smaller (150 observations) than the other three datasets. Since larger datasets reduce problems with multicollinearity a lower tolerance was expected in the continental dataset. Also the correlation of H and Load in the matrix was highest in the continental dataset. However a further exclusion of a predictor variable to account for multicollinearity is not made, because the continental dataset will not be used in the further analysis.

Table 6.36: Multiple stepwise (backwards) regression. H<100$_{cont}$ dataset

N = 150	B	Standard Error B	Signif. Level *p*	Tolerance	Adj. r^2	Standard Error of Estimate	CV %	*p*
Intercept	0.661	0.090	1.4E-11		0.85	0.350	26	<1.0E-17
Thick	-0.411	0.065	2.4E-9	0.94				
H	-0.014	0.004	1.5E-4	0.12				
Load	1.866	0.149	<1.0E-17	0.12				

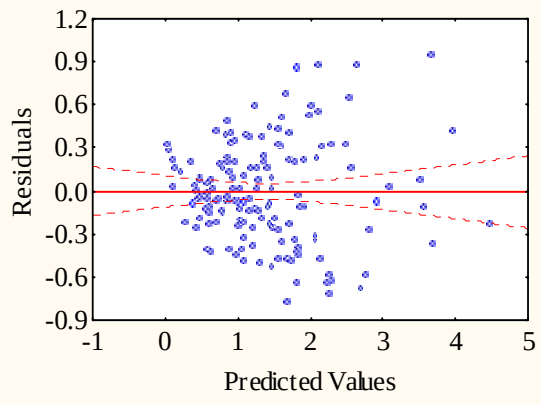

Figure 6.24a: Scatterplot of residuals. Multiple regression. H<100$_{cont}$ dataset.

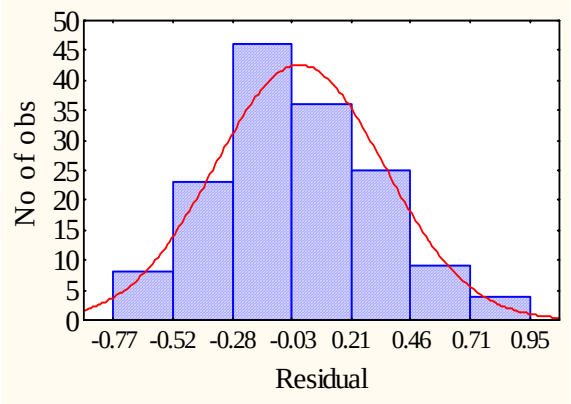

Figure 6.24b: Distribution of residuals. Multiple regression. H<100$_{cont}$ dataset.

The outliers showed the same pattern as in the simple regression analysis and their explanation can be reviewed in Section 6.4.5.

6.4.7 Analysis of the shear strength change rate

In this study, only measured variables at a study plot on a given day were used to forecast the shear strength several days ahead. Unfortunately, this approach does not take real time data into account. Another possible approach is the use of average loading rates or the daily loading rates as measured in weather plots on a daily basis, because load was the most important predictor in the correlation and regression analysis. This procedure was introduced in Section 6.3.

None of the predictor variables showed significant rank correlations with shear strength change rate (Tables 6.23-6.25), and therefore simple regressions were not performed. However, an attempt was made to find a multiple regression equation following the example of Chalmers (2001). He yielded an r^2 value of 0.317 in his multiple stepwise regression analysis (Chalmers, 2001: 73) using a slightly smaller dataset than used in this analysis. Unfortunately, he did not calculate standard errors of the estimate or a coefficient of variation and therefore a direct comparison is not possible.

The correlations with shear strength change rate were only significant for a few variables, and in combination the variables seem to have only a poor fit with the shear strength change rate.

In Table 6.37 the regression results for the Entire dataset are summarized. HS, TG and Strength were picked as predictor variables in the stepwise regression after the removal of fifteen outliers. However, the fit with an r^2 value of 0.12 is very poor. The Kolmogorov-Smirnov test ($D = 0.056$, $p<0.05$) and the Lilliefors test ($p<0.01$) do reject the hypothesis of normality for the residuals (Figure 6.25b). The scatter of the residuals in Figure 6.25a is random and indicates constant variance.

Table 6.37: Multiple stepwise (backwards) regression. Entire dataset. Strength Change Rate.								
N = 661	**B**	**Standard Error B**	**Signif. Level *p***	**Tolerance**	**Adj. r^2**	**Standard Error of Estimate**	**CV %**	***p***
Intercept	0.058	0.01	4.5E-9		0.12	0.081	115	1.0E-17
HS	0.0003	0.00004	1.2E-8	0.76				
TG	0.024	0.005	3.3E-6	0.94				
Strength	-0.016	0.002	8.3E-11	0.80				

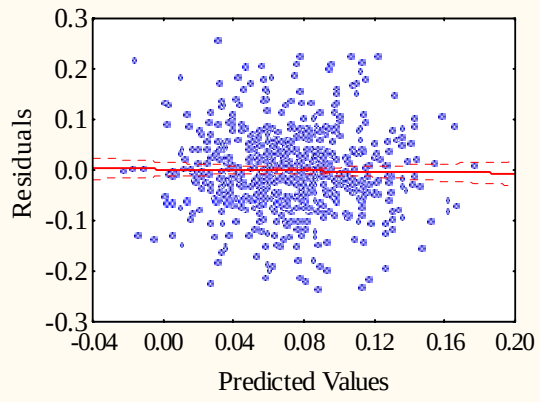

Figure 6.25a: Scatterplot of residuals. Entire dataset. Strength Change Rate.

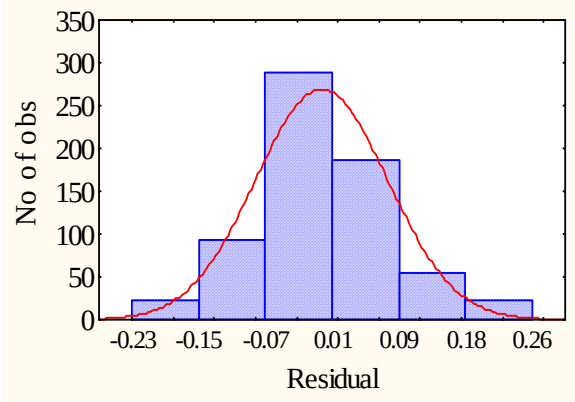

Figure 6.25.b: Distribution of residuals. Entire dataset. Strength Change Rate.

As for the Entire dataset, H was selected in the $H<100_{comb}$ dataset in addition to HS, TG and shear strength (Table 6.38). Eight outliers were removed. The fit of the regression is better (r^2 = 0.21) than in the Entire dataset, but still not satisfactory. The scatter of the residuals is relatively random (Figure 6.26a). Even though the Kolmogorov-Smirnov test (D = 0.051, p<0.15) indicates normality of the residuals, the Lilliefors test (p<0.01) does reject normality of the residuals.

Table 6.38: Multiple stepwise (backwards) regression. H <100 dataset. Strength Change.

N = 526	B	Standard Error B	Signif. Level *p*	Tolerance	Adj. r^2	Standard Error of Estimate	CV %	*p*
Intercept	0.041	0.011	1.7E-4					
HS	0.0002	0.0001	1.3E-5	0.83				
H	0.001	0.0002	3.7E-8	0.46	0.21	0.073	102	<1.0E-17
TG	0.023	0.005	1.1E-6	0.95				
Strength	-0.042	0.005	1.1E-15	0.49				

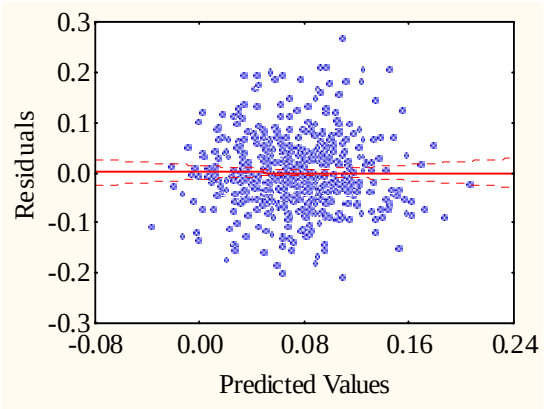

Figure 6.26a: Scatterplot of residuals. H<100 dataset. Strength Change Rate.

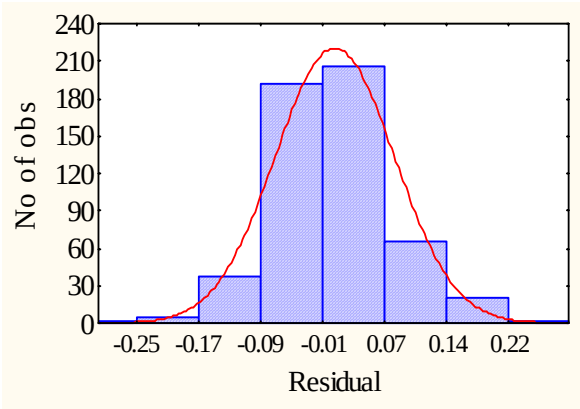

Figure 6.26b: Distribution of residuals. H<100 dataset. Strength Change Rate.

The $H<100_{inter}$ dataset yielded an r^2 value of 0.21, again a poor fit. Age, Load, TG and SlabDens were chosen as predictor variables after the removal of five outliers (Table 6.39). The scatter shows, as already observed in the other datasets, a relatively random pattern (Figure 6.27a). Normality of the residuals was not rejected with the Kolmogorov-Smirnov test (D = 0.051, p<0.15), but was rejected with the Lilliefors test (p = 0.01) (Figure 6.27b).

Table 6.39: Multiple stepwise (backwards) regression. H<100$_{inter}$ dataset. Strength Change.								
N = 357	**B**	**Standard Error B**	**Signif. Level *p***	**Tolerance**	**Adj. r^2**	**Standard Error of Estimate**	**CV %**	***p***
Intercept	0.177	0.015	<1.0E-17		0.21	0.07	82	3.0E-17
Age	-0.002	0.0005	3.3E-4	0.41				
Load	0.053	0.013	2.9E-5	0.31				
TG	0.031	0.006	3.8E-7	0.97				
SlabDens	-0.0006	0.0001	6.0E-6	0.36				

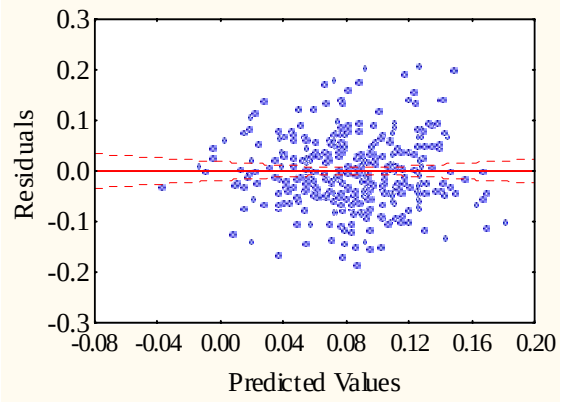

Figure 6.27a: Scatterplot of residuals. H<100$_{inter}$ dataset. Strength Change Rate.

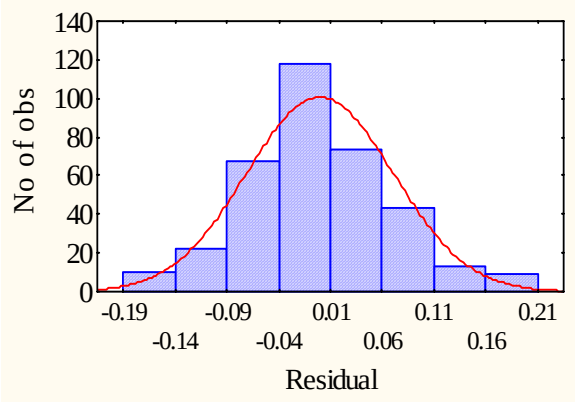

Figure 6.27b: Distribution of residuals. H<100$_{inter}$ dataset. Strength Change Rate.

Only two predictors, H and Strength, were selected by the stepwise regression analysis in the dataset from the continental snowpack (Table 6.40). However, the intercept did not have a significant B value. The fit after removing 5 outliers was 0.17, the scatter is random (Figure 6.28b), and normality was not rejected with the Kolmogorov-Smirnov test ($D = 0.096$, $p < 0.1$), but was rejected with the Lilliefors test ($p < 0.01$) (Figure 6.28b).

Table 6.40: Multiple stepwise (backwards) regression. H<100$_{cont}$. Strength Change Rate.								
N = 162	**B**	**Standard Error B**	**Signif. Level *p***	**Tolerance**	**Adj. r^2**	**Standard Error of Estimate**	**CV %**	***p***
Intercept	0.014	0.015	0.325		0.17	0.074	164	4.5E-7
H	0.002	0.001	7.2E-7	0.35				
Strength	-0.062	0.011	1.2E-7	0.35				

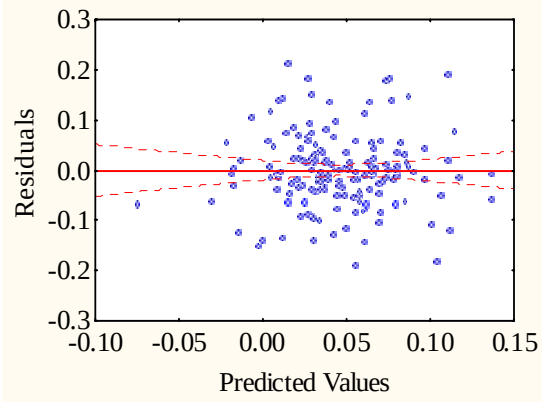

Figure 6.28a: Scatterplot of residuals. H<100$_{cont}$ dataset. Strength Change Rate.

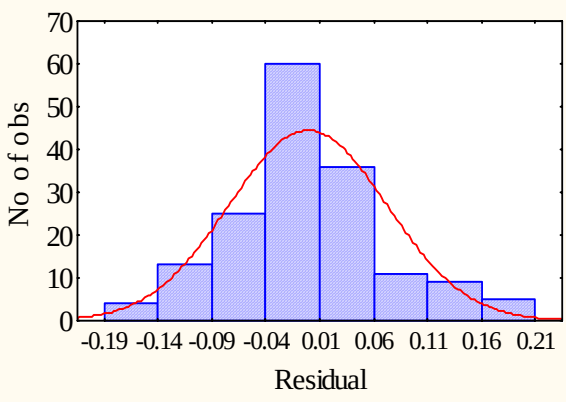

Figure 6.28b: Distribution of residuals. H<100$_{cont}$ dataset. Strength Change Rate.

None of the multiple regression models are promising predictor models for the strength change rate based on their r^2 values. Chalmers' (2001) model to predict the strength change rate was with an r^2 value of 0.317, only slightly better. However, his result could not be repeated and the average and daily loading rates will be analyzed along with the multiple models found.

6.4.8 Model selection

Table 6.41 summarizes the regression results for the shear strength analysis and is the basis for the model selection.

Table 6.41: Summary of regression results for strength.						
Regression		r^2	Significant B and intercept	Variance of residuals	Normal distribution of residuals	Tolerance accepted
Simple logarithmic	Entire	0.86	Yes	constant	Yes	-
	$H<100_{comb}$	0.82	Yes	constant	Yes	-
	$H<100_{inter}$	0.82	Yes	constant	No	-
	$H<100_{cont}$	0.82	Yes	No	Yes	-
Multiple	Entire	0.88	No	No	No	Yes
	$H<100_{comb}$	0.83	Yes	o.k.	Yes	Yes
	$H<100_{inter}$	0.84	No	No	Yes	Yes
	$H<100_{cont}$	0.85	Yes	No	Yes	No

The results for both, the simple logarithmic and multiple regressions, yielded good r^2 values of 0.82 to 0.88. However, the following regressions will be excluded from further analysis based on physical reasons, the significance of the B values of the regression coefficients and the intercept, the distribution and variance of the residuals and the r^2 values:

1. the regressions solely based on data from a continental snowpack ($H<100_{cont}$ logarithmic and multiple), because they were only used to assess whether the data are similar to the one from an intermountain snow climate so that it is justified to use the combined dataset and the low tolerance (< 0.2)
2. the regressions from the entire dataset (logarithmic and multiple), because the H<100 datasets are more relevant for skier triggering and therefore for this thesis. In addition the multiple regression the intercept was insignificant, the residuals were not normally distributed and the variance of the residuals not constant
3. the logarithmic regression for the intermountain dataset ($H<100_{int}$), because the residuals were not normally distributed
4. the multiple regression from the intermountain dataset ($H<100_{int}$), because the intercept of the regression was insignificant and the variance of the residuals not constant

Consequently the power law in Equation 6.7b and the results from the multiple regression in Table 6.38 will be further analyzed. The multiple regression is re-written as

$$\Sigma = 0.642 - 0.265\ \text{Thick} + 0.044\ \text{T-5} + 0.001\ \text{HS} - 0.009\ \text{H} + 1.573\ \text{Load}$$

Equation 6.8

The Forecasting Model (Equation 6.1) for the $H<100_{comb}$ dataset for the power law using average loading rates for the intermountain snow climate used in further analysis is re-written as

$$\Sigma_j^* = 1.51\ \text{Load}^{0.93} + 0.062\ \Delta t_{ij} \qquad \textit{Equation 6.9a}$$

and for the multiple regression as

$$\Sigma_j^* = (0.642 - 0.265\ \text{Thick} + 0.044\ \text{T-5} + 0.001\ \text{HS} - 0.009\ \text{H} + 1.573\ \text{Load}) + 0.062\ \Delta t_{ij}$$

Equation 6.9b

The Forecasting Model (Equation 6.1) for the $H<100_{comb}$ dataset for the power law using daily loading rates is re-written as

$$\Sigma_j^* = 1.51\ \text{Load}^{0.93} + \text{Pcp}_{ij} \qquad \textit{Equation 6.10a}$$

and for the multiple regression as

$$\Sigma_j^* = (0.642 - 0.265\ \text{Thick} + 0.044\ \text{T-5} + 0.001\ \text{HS} - 0.009\ \text{H} + 1.573\ \text{Load}) + \text{Pcp}_{ij}$$

Equation 6.10b

Table 6.42 summarizes the coefficient of determination and the absolute error between the estimated and measured shear strength. As in the analysis of faceted layers the start of the interval refers to the estimated shear strength calculated on days with snowpack observations (Equations 6.7b and 6.8) and the end of the interval refers to the calculated shear strength with the Forecasting Model (Equations 6.9a and b for average loading rates and Equation 6.10a and b for daily loading rates). All of the regressions from the strength change rate are poor and are omitted from further analysis. Instead the average and the daily loading rates will be used to forecast the strength up to eight days ahead.

Table 6.42: Fit of Forecasting model to data used to build the model for surface hoar. H<100comb dataset

	Power law model			Multiple regression model		
Interval	Start	End (avg)	End (daily)	Start	End (avg)	End (daily)
N	579	482	286	492	414	255
r^2	0.72	0.66	0.62	0.78	0.71	0.63
Mean deviation (kPa)	0.421	0.440	0.473	0.324	0.434	0.456

The multiple regression for the combined dataset will be used in the remaining analysis of this study because:

1. the multiple regression model yielded better r^2 values at the start and the end of each interval (Table 6.42)
2. the multiple selection of predictor variables is physically more explicable, because shear strength does not only depend on Load

Equation 6.8 and 6.9 will be tested on two time series in the next section.

6.4.9 Model testing

The first test series is from an intermountain snow climate and was buried on the 10th of February 1997 at Mt. St. Anne (Appendix 2). This time series comprises nine measurement days over a period of 45 days. In Figure 6.29 The estimated shear strength at

the start of an interval (Equation 6.8) and the shear strength forecast at the end of each interval for the calculations average and daily loading rates (Equations 6.9b and 6.10b) are plotted along with the measured values.

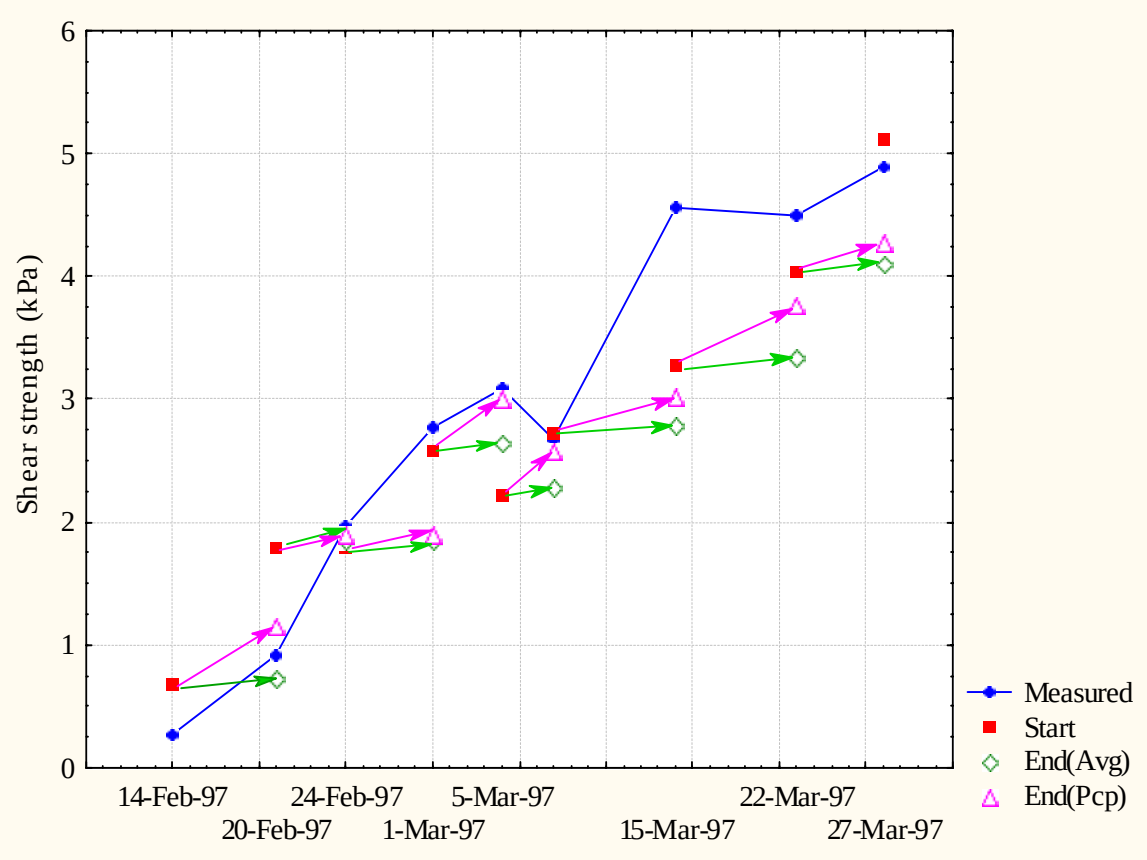

Figure 6.29: Model testing: forecasting shear strength at Mt.St. Anne (surface hoar layer 10 February 1997). Average (green arrows) and daily loading rates (pink arrows).

At the start of the interval the model explained 85% of the variability between the observed and predicted values. The Forecasting Model performed with a fit of 88% and 87% using the average and the daily loading rates respectively. The mean average error at the start of the interval was 0.506 kPa, at the end of the interval for the average loading rate 0.722 kPa and at the end of the interval for the daily loading rate 0.530 kPa. Figure 6.28 shows that the model using daily loading rates predicts the shear strength change over one measurement interval more realistically than the model based on average loading rate, even

though the correlation is slightly weaker. However the absolute average error was lower by 0.192 kPa for the model using the daily loading rates. The daily loading rates reflect better the trend of each measurement interval. On most of the measurement days the models at the start and the end of the intervals underestimated the shear strength. On day 4 and 10 the model overestimates the strength at the start of the interval by 150% and 96% respectively, which might be because pulling shear frames is difficult when the weak layer is still shallow and the slab is relatively uncohesive. After day ten the model underestimates the shear strength by an average error of 14%. The shear strength was underestimated consistently using Equation 6.10a; the average error equalled 21%. Using Equation 6.10b underestimated the shear strength on all but day10. The underestimation equalled 15%. Day ten was overestimated by 27% (0.245 kPa). On day 33 all three models underestimated the shear strength with 38%, 39% and 34% at the start, the end using average and the end using daily loading rates respectively. This can be due to spatial variability.

The second test series was measured on the surface hoar layer buried on the 28th of January 2001 at Mt. Fidelity and consists of 15 observations over 58 days (Appendix 3). In Figure 6.31 the estimated shear strength at the start of an interval (Equation 6.8) and the shear strength forecast at the end of each interval (Equation 6.9b and 6.10b) are plotted along with measured values. The underestimation on day 21 is likely because of a measurement error in the thickness of the weak layers. The thickness was reported with 3 cm whereas the layer in the rest of the series was measured with a maximum of 1.2 and a minimum of 0.3. Ignoring this value the average error equalled 14%, 21%, 19% (start, avg, pcp) respectively. At the end of the series the underestimation is greater than at the start and corresponds to layers which are deeper buried than 100 cm and these deeper layers were not used in the model building process. The models explained 90%, 96% and 94% of the variability between the observed and predicted values at the start and the end (average and daily loading rates) respectively. The average absolute errors were 0.377, 0.515 and 0.436 kPa. Similar to the first test series the r^2 value showed a better fit when using the average loading rates, but the average absolute was higher compared to the model using the daily loading rates. However the advantage of daily loading rates is not as apparent, but of

advantage for deeper burials and daily loading would be better for periods of above average or below average loading, and these are important for forecasting.

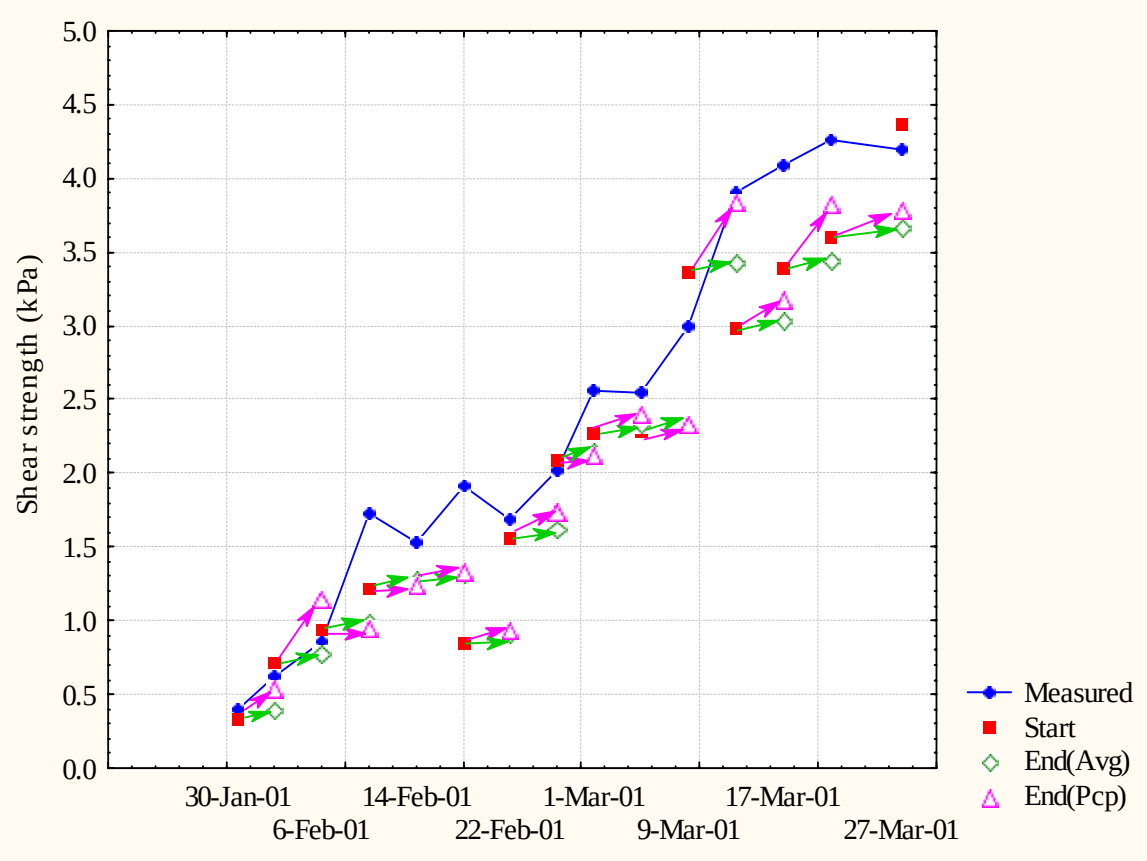

Figure 6.31: Model testing: forecasting shear strength at Mt. Fidelity (surface hoar layer 28 January 2001). Average (green arrows) and daily loading rates (pink arrows).

Even though the fit of the model using average loading values was reasonable, the model with daily loading rates fitted the test series better and consequently only this value will be used in the further calculations in this thesis.

6.5 Summary

Stability indices are calculated using the shear strength of buried weak layers, which are potential failure layers for slab avalanches. However, shear strength measurements with a shear frame are not standard observations in forecasting operations. In order to be able to

incorporate stability indices into a forecasting model daily shear strength values have to be calculated. In Section 6.2 an Interval Model was introduced to forecast the shear strength of layers of faceted crystals and surface hoar, which requires that two formulas be developed, one to estimate the shear strength on a day with manual snowpack observations and one to determine the shear strength change rate between snowpack observations. Simple, multiple and logarithmic regression analyses were performed in order to find the best models. In the analysis of layers of faceted crystals (Section 6.3) a power law relationship of shear strength and load yielded the best performance. The shear strength change rate is based on long-term average loading rates for the Columbia Mountains and daily loading rates, because the shear strength change rate did not significantly correlate with most of the predictor variables. Also under consideration was the burial depth of faceted layers. For this the entire dataset was considered but also the part of the dataset where the depth of the weak layer did not exceed 170 cm in order to account better for skier-triggered avalanches. However the best fit (r^2) was achieved when using the entire dataset with 77%-83% compared to 58%-63% when using the average and daily loading rates. The Forecasting Model using daily loading rates and the combined dataset were in the end chosen and are used in the calculation for stability indices in the remainder of this thesis.

The dataset of surface hoar is larger than the one for facets and a more refined model could be found (Section 6.4). In order to predict skier-triggered dry slab avalanches in the Columbia Mountains four datasets were considered. The first dataset contained all the observations available and the other three considered only measurements when the weak layers was not deeper than 100 cm. Additionally shear frame measurements from the continental snowpack (Rocky Mountains) and the intermountain snowpack (Columbia Mountains) were separated to assess whether the shear strength of weak layers exhibited different behavior in the different snow climates. The coefficient of determination of a simple logarithmic and multiple regression yielded good results with values of 0.82 to 0.88, however the assumption of the normality of the residuals and the constancy of variance were violated in some of the datasets. A distinction between the continental and the intermountain snowpack data seemed not to improve the fit. In the end the multiple regression model based on data with less than 100 cm depth and from both snowpacks was

chosen because it accounts for the relevant depth for skier triggering and the shear strength depends on more than load. The fit yielded an r^2 of 0.78. A multivariate model for the shear strength change rate could not be found and again the average and daily loading rates were used in the Forecast Model. The r^2 was 0.71 and 0.63 respectively. However the test series showed that daily loading rates have greater value because specific weather patterns are better reflected in the calculations and daily loading rate accounts better periods with below or above average snowfall.

7. Forecasting skier-triggered avalanches on persistent weak layers

7.1 Introduction

In Chapter 6 empirical models were developed that could forecast the shear strength of persistent weak layers in order to be able to calculate daily stability indices and incorporate these into avalanche forecasting models. Nearest Neighbour models for avalanche forecasting have made little use of snowpack properties; however, slab thickness *(H),* slab load (Load) and a skier stability index (Sk38) have proven useful for regional avalanche forecasting in the Columbia Mountains of western Canada (Zeidler and Jamieson, 2004). In this chapter, 26 meteorological, snowpack and elaborated variables including Sk38, *H* and Load will be explored in regard to their predictive value for avalanche forecasting. A daily skier instability index (DSI) is developed as a response variable using skier-triggered avalanche activity on persistent weak layers and stability ratings at the end of the day. The aims are:

1) to assess the predictive potential of the snowpack properties Sk38, *H* and Load using rank correlation and classification tree analysis, and
2) to incorporate Sk38, *H* and Load into a Nearest Neighbour model and verify that these stability measures improve the performance of a Nearest Neighbour model for skier-triggered avalanches.

The focus is to forecast better skier-triggered dry slab avalanches on persistent weak layers and does not take into account loose snow avalanches, wet avalanches or dry slab avalanches on weak layers of precipitation or decomposed and fragmented precipitation particles (non-persistent weak layers). However, dry slab avalanches on non-persistent weak layers are analyzed in Chapter 8.

7.2 Methods

7.2.1 Dataset selection

Time periods during which a persistent weak layer was present in the study plot and tested with the shear frame were extracted from the entire datasets of two forecasting areas, Blue River and Rogers Pass/Adamants. Typically, testing started when a persistent weak layer had been buried for two to five days and continued until the weak layer had been buried for approximately six weeks. The following were excluded from this study:

- Burial depths of surface hoar layers more than 130 cm on a measurement day, because the extrapolation of shear strength in Chapter 6 is based on slab thickness and because skiers are less likely to trigger avalanches when the slab thickness is greater than 100 cm.
- Observations where the shear strength of faceted layers significantly decreased over one measurement interval, since such apparent decreases are probably caused by spatial variability in the study plot or measurement errors.
- Days on which Sk38 could not be calculated because the ski penetration exceeded the depth of the weak layer as calculated in Equation 2.8.
- Observations when the depth of the persistent weak layer was less than 30 cm (Schweizer et al., 2003), and where the depth was less than 40 cm and the load less than 0.4 kPa. This was done to exclude days with possible avalanches that were not likely to threaten a skier, because many such slabs are not sufficiently cohesive for fracture propagation. Also, small avalanches were not consistently recorded. It is expected that shallower slabs may be better forecast with the model for non-persistent weak layers.

The testing of the persistent weak layers did not always begin when avalanche activity on this layer was first reported. On days between the burial date and the first measurement, Load was calculated by subtracting the daily loading rate from the first measured value and H was estimated from the daily measurements of the height of new snow and average settlement rates for the Columbia Mountains.

In the case of more than one persistent layer in the study plot, the weakest one was chosen because the shear strength of a weak layer correlated strongly with its stability (Föhn and Camponovo, 1997). However, only weak layers for which a positive Sk38 could be calculated were considered.

7.2.2 Response variable

A daily skier instability index, DSI, was developed. DSI was assigned a value of one for each day in which one or more skier-triggered, dry slab avalanches on a persistent weak layer were reported, or the ski guides operating in the area rated the stability at the end of the day as fair-poor, poor or very poor (Canadian Avalanche Association, 2002: 74), and DSI = 0 for the other days. This index allows the inclusion of days on which skier-triggered avalanches were likely, but due to limited conditions for flying (and hence helicopter skiing) or the terrain selection of the guides, no dry slab avalanches on persistent weak layers were skier-triggered.

7.2.3 Predictor variables

The explanation and calculation for Sk38 were introduced in Section 2.4.3, the snowpack variables Load and *H* need further explanation:

Load: Load is the snow density integrated over slab thickness. The shear strength of weak layers increases with slab load (Jamieson et al., 2001) but lags of a few days are likely for persistent weak layers (Chalmers, 2001: 79-84). In this study, load is calculated as follows:

$$Load_i = Load_{i-1} + PcpYLoad \qquad \textit{Equation 7.1}$$

where $Load_i$ is the slab load on the day to be forecast, $Load_{i-1}$ is the load on the previous day and PcpYLoad is the load in kPa over the last 24 hours as measured with a precipitation gauge.

Zeidler and Jamieson (2002) used an average loading rate of 0.062 kPa d^{-1} for tree-line elevation in the area around Blue River although it is expected that the daily loading rates are more accurate. Calculating Sk38 required shear strength values of weak snowpack layers. On days with snowpack observations, these were measured using a shear frame (Jamieson and Johnston, 2001). On days without manual snowpack observations, daily shear strength values were calculated by applying the empirical model to forecast the shear strength of surface hoar layers (Section 6.4) and the model to forecast the shear strength of faceted layers (Section 6.3).

H: Slab thickness is an important variable to be considered when looking at skier triggering. The shear stress induced by skiers on a weak layer decreases with increasing depth of the weak layer. Consequently more deeply buried weak layers are harder to trigger than shallower weak layers. However layers shallower than 30 cm release hazardous slabs less often than deeper weak layers. Slab thickness (Equation 7.2) was predicted on days with no snowpack observations:

$$H_i = 0.95\, H_{i-1} + \mathrm{HNY} \qquad \textit{Equation 7.2}$$

where H_i is the slab thickness on the day to be forecast, H_{i-1} is the slab thickness on the previous day, and 0.95 is an average daily settlement factor for the Columbia Mountains (Zeidler and Jamieson, 2004).

The average value of settlement of slabs, rather than measured values was used because it is available in the morning when route selection decisions are being made, while measured values would not be available until hours later when guides or technicians might visit the study plot.

Daily values of Sk38 can be calculated with values of the shear strength of the weakest persistent snowpack layer, overlying load and slab thickness. This allows assessment of the predictive value of Sk38, *H* and Load and any combination of these three predictors to forecast DSI.

All variables considered in this study are listed in Table 7.1.

Table 7.1 Daily predictor variables

Abb.	Definition	Blue River	Rogers Pass	Adamants
T	Air temperature at a.m. (°C)	✓	✓	✓
RH	Relative humidity at a.m. (%)	✓	✓	~
WS	Wind speed at a.m. (km h^{-1})	✓	✓	~
WrunY	24 h wind run for previous day (km)	✓	✓	~
HNY	Height of new snowfall for previous 24 h (m)	✓	✓	✓
HS	Height of snowpack at a.m. (m)	✓	✓	✓
dT	Change in a.m. air temperature from previous day (°C)	✓	✓	~
PcpY	Water equivalent of precipitation on previous day (mm)	✓	✓	~
Strm	Cumulative new snowfall (storm) since last day with less than 0.3 mm of precipitation (m)	✓	✓	~
TminY	Minimum temperature for previous day (°C)	✓	✓	✓
TmaxY	Maximum temperature for previous day (°C)	✓	✓	✓
RHmnY	Minimum relative humidity for previous day (%)	✓	✓	~
RHmxY	Maximum relative humidity for previous day (%)	✓	✓	~
WSa	Average upper air wind speed (km h^{-1})	✓	~	~
Baro	Barometric pressure at a.m. (mbar)	~	~	✓
dBaro	Change in barometric pressure in past three hours. 1 = steady, 2 = rising, 3 = falling, 4 = rising rapidly, 5 = falling rapidly (CAA, 2002: 10)	~	~	✓
Sky	1 = clear, 2 = scattered, 3 = broken, 4 = overcast, 5 = obscured (CAA, 2002: 2)	~	~	✓
Pcp	1 = light rain, 2 = nil, 3 = snow < 1 cm/h, 4 = snow ~ 1 cm/h, 5 = snow ~ 2 cm/h, 6 = snow ~ 3 cm/h (CAA, 2002: 3)	~	~	✓
WSad	1 = calm, 2 = light, 3 = moderate, 4 = strong, 5 = extreme (CAA, 2002: 9)	~	~	✓
HSTD	Height of storm snow since last a.m. reading without snow	~	~	✓
HSTM	Height of storm snow since last standard observation	~	~	✓
NaPrev	Number of natural avalanches on previous day	✓	~	~
TriPrev	Number of skier-triggered dry slab avalanches on previous day	✓	~	✓
Sk38	Skier stability index	✓	✓	~
Load	The weight of overlying slab above a weak layer (kPa)	✓	✓	~
H	Slab thickness above a weak layer (cm)	✓	✓	~

7.2.4 Analytical methods

First, two methods that do not assume the distribution or continuity of the variables were used, because the response variable DSI has only two values (0 or 1) (Jarret and Kraft, 1989: 600):

Spearman rank correlations indicate the degree of association of each predictor variable with the response variable (Johnson and Bhattacharyya, 1996: 632).

Classification trees are used to determine the importance of predictor variables associated with DSI when meteorological and snowpack variables are used in combination, and to understand the interactions. Classification trees were constructed using all variables in Table 7.1, including or excluding the snowpack variables Sk38, Load and *H*.

Finally, the computer program Cornice was used to assess the predictive merit of snowpack properties for avalanche forecasting with a Nearest Neighbour (NN) model. In addition, the suitability of NN models for computer assisted avalanche forecasting in backcountry operations was assessed.

7.3 Blue River

The first study area lies in the Columbia Mountains of western Canada near Blue River, British Columbia, BC (see Figure 3.1, Chapter 3). Meteorological variables were available from a remote weather station at Mt. St. Anne at an altitude of 1900 m. Snowpack data and shear strength values were recorded once or twice a week in a level study plot within 300 m of the meteorological station. Stability indices calculations were possible for nine winters from 1995/1996 to 2003/2004. The last two years were used as a test sample whereas the first seven years were used for the model development and are referred to as the learning sample. Snowpack parameters such as slab load (Load) and slab thickness (*H*) as well as the stability index Sk38 were extrapolated for days without manual snowpack observations. Avalanche activity data and stability evaluations were provided by the heli-skiing operation around Blue River. The study area is approximately 5000 km^2, while most skiing during the winter takes place within 30 km of the study plot and the weather station (Jones and Jamieson, 2001).

Table 7.2 summarizes the available time series for which extrapolations of shear strength were possible and the number of days per series after the exclusions mentioned in Section 7.2.1 were applied.

Table 7.2: Time series measurements of persistent weak layers used in Blue River dataset. Bolded italics indicate the test samples.

Weak Layer date	Crystal form	No. of days, N
17-Feb-1996	Surface hoar	39
22-Nov-1996	Facets	41
17-Jan-1997	Surface hoar	28
10-Feb-1997	Surface hoar	35
8-Dec-1997	Surface hoar	30
13-Feb-1998	Surface hoar	19
25-Feb-1998	Surface hoar	23
24-Dec-1998	Facets	17
30-Dec-1999	Surface hoar	28
31-Jan-2000	Surface hoar	29
7-Dec-2000	Surface hoar	45
20-Jan-2001	Surface hoar	30
23-Feb-2001	Surface hoar	3
8-Mar-2001	Surface hoar	23
1-Jan-2002	Surface hoar	13
7-Jan-2002	Facets	27
16-Feb-2002	Surface hoar	37
5-Apr-2002	Facets	14
25-Dec-2002	***Surface hoar***	***29***
19-Jan-2003	***Surface hoar***	***20***
15-Feb-2003	***Surface hoar***	***40***
3-Mar-2004	***Surface hoar***	***12***
9-Mar-2004	***Facets***	***14***

In total, the dataset consisted of 596 days in 23 time series. As described in Section 7.2.2 a skier instability index DSI was defined and Table 7.3 gives an overview of the dataset with the number of days defined as DSI = 1 and DSI = 0.

Table 7.3: Summary of dataset: Blue River.					
	N	DSI (days) 1	0	Facets series (days)	Surface hoar series (days)
Learning sample	481	163	318	4 (99)	14 (382)
Test sample	115	64	51	1 (14)	4 (101)
Total	596	227	369	5 (113)	18 (483)

Zeidler and Jamieson (2004) analyzed a similar dataset from Blue River, but since then significant changes have been made to the dataset to improve the forecast of the proposed model. The changes include:

- shear strength adjustments to the normal load effect during shear frame measurements
- calculation of Load using daily precipitation measures rather than the average loading rate for the Columbia Mountains
- cut-off for surface hoar extrapolations at 130 cm instead of a time related cut-off at 30 days
- update of the DSI index by using an additional source of avalanche occurrences
- exclusion of days when the weak layer depth is less than 30 cm or less than 40 cm when the load was less than 0.4 kPa
- missing values of the upper wind speed were updated by data from the forecasting service in Kelowna, BC.

In the earlier study, the data from the last two winters were not available and a test sample was not used. Consequently the learning sample in the new dataset resembles the dataset of the earlier study. However due to the changes made in regard to the shear strength extrapolations and updating of predictor variables, the two datasets are different. In the previous study a total of 515 days were considered with 128 defined unstable days (DSI = 1) whereas the learning sample of the updated dataset consists of 481 days with 163 defined unstable days, which means that the latter dataset is more balanced. Skier-triggered avalanches on non-persistent weak layers or low stability ratings due to storm snow instabilities were defined as stable days (DSI = 0). The differences in the datasets were

likely to influence the outcome of the analysis. Differences in the results will be discussed as they emerge.

7.3.1 Spearman rank correlations

In Table 7.4 the Spearman rank correlation results with DSI are displayed.

Table 7.4. Spearman rank correlations with DSI. Insignificant variables (p > 0.05) are in italics.

Variable	Valid N	Spearman R	p-level
Load	596	-0.330	1.3E-16
SK38	596	-0.320	1.2E-15
TriPrev	596	0.243	1.9E-9
HNY	595	0.221	5.5E-8
PcpY	595	0.200	8.5E-7
Strm	596	0.198	1.1E-6
Wsa	595	0.198	1.1E-6
RHmxY	595	0.196	1.5E-6
RHmnY	596	0.180	1.0E-5
RH5	596	0.153	1.8E-4
HS5	596	-0.150	2.5E-4
H	596	-0.116	0.005
WrunY	582	0.117	0.005
NaPrev	*596*	*0.051*	*0.213*
T5	*596*	*-0.045*	*0.276*
WS5	*585*	*0.037*	*0.370*
TmaxY	*596*	*-0.026*	*0.528*
dT5	*596*	*0.021*	*0.615*
TminY	*596*	*0.011*	*0.789*

Significant correlations with DSI are:

Positive Correlations

- **Skier-triggered dry slab avalanches on previous day (TriPrev):** Avalanches often occur on consecutive days because snowfall (loading) due to a snow storm often continues for several days and because persistent weak layers require time to adjust to increased overlying load (Chalmers, 2001: 79-84).

- **Precipitation:** Higher values of precipitation (PcpY), new snow (HNY) and storm snow (Strm) accumulations indicate recent loading on the weak layers and the positive correlation suggests that the higher the precipitation values, the more likely that there will be unstable days.
- **Wind:** When wind speed (WSa) and the wind run (Wrun) are greater, more snow is transported, loading the weak layer faster and thus the formation of wind slabs is promoted (McClung and Schaerer, 1993, p. 157-158).
- **Humidity**: Higher values of RH5, RHmnY and RHmxY are associated with stormy weather (precipitation) and slab formation (McClung and Schaerer, 1993: 161) and consequently with avalanche activity and therefore instability.

Negative Correlations

- ***H*:** Having a negative correlation implies that a thinner slab thickness is associated with skier-triggered dry slab avalanches. Skier stress decreases with increasing depth (Föhn, 1987a; Schweizer and Camponovo, 2001). Consequently deeper weak layers are less often triggered by skiers. Föhn (1987a) found that since the calculated static skier stress at 1 m is only 10% of the stress acting on the slab due to gravity, skiers are not efficient triggers where the slab is more than a meter thick. In addition, Johnson (2000: 57) found a positive correlation of slab thickness with the shear strength of faceted layers, implying that deeply buried weak layers of faceted crystals are usually stronger.
- **Load:** The negative correlation can be interpreted in the same way as for slab thickness. Load is the primary snowpack variable that affects the shear strength of layers of faceted crystals and surface hoar layers (Chalmers, 2001: 52). Thicker slabs typically overlie stronger facet layers (Johnson, 2000: 57; Zeidler and Jamieson, 2002). Additional slab load causes increased densification (Kojima, 1967; Conway and Wilbour, 1999) and increased bonding.

- **Sk38:** This result was expected since lower values of Sk38 indicate lower stability and increased probability of skier triggering (Jamieson, 1995: 148-158, 215-221; Jamieson and Johnston, 1998).
- **HS:** Winters with a deeper snowpack typically have less clear weather and consequently the surface hoar layers that form generally consist of smaller crystals, which gain shear strength faster (Jamieson and Johnston, 1998). Also, the slabs that bury those persistent weak layers are likely to thicken sooner. Consequently, the stress induced by skiers less frequently penetrates deeply enough to cause the weak layer to fail. When the snowpack is relatively thin, densification takes longer and weak layers are preserved for a longer time.

Even though temperature and wind are thought to be important factors in avalanche forecasting, the correlations were not significant in this dataset. For temperature this may be because the focus was on dry slab avalanches from December to March, or because only daily temperature variables were considered when hourly changes might have been more relevant. The non-significant correlations for ridge-top winds (WS5) may be due to limited wind effects in the North Columbia Mountains.

Compared to the dataset used in Zeidler and Jamieson (2004), Load changed positions with Sk38 and was ranked as the highest correlated variable with DSI. This is possibly because Load is calculated using the daily loading rate rather than the average loading rate as was done in the previous study. The slab thickness *H* was ranked fourth in the old dataset and is now in twelfth position possibly because days with a burial depth of less than 30 cm were excluded. WSa is more significant in the updated dataset, likely because the number of missing values was reduced. Most of the non-significant variables stayed that way except the Wrun, which is now slightly significant, perhaps because more days are used.

7.3.2 Classification tree analysis

Multivariate statistical methods were applied in this analysis to assess the association of snowpack properties, especially Sk38, Load and *H*, with DSI.

Table 7.5 shows the results for the global cross-validation for the learning and test sample for several possible input parameter combinations using snowpack properties.

Table 7.5. Global cross-validation results from classification tree. The learning sample (L) consists of 142 unstable days and 323 stable days. The test sample (T) consists of 60 unstable days and 52 stable days.

		Misclassification						
		Unstable days (DSI = 1)		**Stable days (DSI = 0)**				**Priors DSI=0/**
Variables	**Sample**	**No**	**(%)**	**No**	**(%)**	**Cost**	**Error**	**DSI=1**
Including Sk38, Load and *H*	L	50	(35)	134	(42)	0.378	0.026	0.41/
	T	16	(27)	30	(58)	0.394	0.044	0.59
Excluding Sk38, Load and *H*	L	48	(34)	141	(44)	0.379	0.026	0.41/
	T	22	(37)	23	(44)	0.398	0.046	0.59
Including SK38	L	54	(38)	121	(38)	0.378	0.027	0.41/
	T	23	(38)	13	(25)	0.329	0.045	0.59
Including *H*	L	45	(32)	166	(51)	0.398	0.026	0.41/
	T	23	(38)	23	(44)	0.408	0.047	0.59
Including Load	L	50	(35)	127	(39)	0.369	0.026	0.42/
	T	14	(23)	21	(40)	0.305	0.043	0.58
Including SK38 + *H*	L	52	(37)	127	(39)	0.377	0.026	0.41/
	T	26	(43)	12	(23)	0.350	0.045	0.59
Including SK38 + Load	L	48	(34)	137	(42)	0.374	0.026	0.42/
	T	14	(23)	18	(35)	0.281	0.042	0.58
Including *H* + Load	L	49	(35)	136	(42)	0.377	0.026	0.42/
	T	19	(32)	18	(35)	0.329	0.045	0.58

Comparing the results from the global cross-validation of each classification tree the results are quite similar. However, when choosing the best tree model, unstable days should be forecast more accurately, although the forecast for stable days should also not be too inaccurate because this would lead to the forecasting of a number of unstable days on which no avalanches are actually skier-triggered (false alarms) and as a result, the trust in the forecast model may diminish. The overall misclassifications, considering the set priors, are expressed in the misclassification costs. The lowest cost in the learning sample (test

sample) was calculated when including Load with 0.369 (0.305) and including Sk38 and Load with 0.374 (0.281). Looking at the misclassified cases reveals that including Sk38 and Load improved the classification by two cases for the unstable days, but worsened the classification of stable days by ten cases in the learning sample. In the test sample the misclassified unstable days were equal, though the misclassification with Load and Sk38 improved by three cases. It seems that by using Sk38 in addition to Load, unstable days can be more accurately classified.

A comparison including and excluding all three snowpack properties (Sk38, *H* and Load) showed that the learning sample better classified unstable days when excluding the three predictors, with 48 days compared to 50 days misclassified. A better classification of stable days was obtained by including the three predictors (134 cases compared to 141). However, the costs are similar, with 0.378 when including the three properties and 0.379 when excluding them. The test sample however had a lower cost of 0.394 when including the three predictors compared to 0.398 when they were excluded. This result is different from the results in the old dataset where both the classification of stable and unstable days was improved when using Sk38, Load and *H* (Zeidler and Jamieson, 2004), possibly because of the update of the DSI index.

Including *H* has the highest misclassification cost in both the learning and test sample with 0.398 and 0.408, respectively. However, the misclassified unstable days were lowest with 45 of out 142 days, though it predicts the stable days with only 51% accuracy.

Including Sk38 seems to predict unstable days least accurately with 54 misclassified cases in the learning sample, but provides the best classification for stable days. It was expected that this would be the reverse. In combination with Load however, the misclassification of unstable days is reduced by six cases. Incorporating only one of the three analysed predictors showed that Load has the highest predictive potential. Sk38 best predicts stable days, but the misclassification for unstable days is higher with 38%. *H* better predicted unstable days with 32%, however the misclassification of stable days was 51%.

The further analysis of Load and *H* in combination is rejected because the misclassification cost is higher compared to using Load alone and it is assumed that *H* does not add significant predictive value to the model. Also inclusion of Sk38 and *H* separately

will not be further analysed, because in combination, they have a lower misclassification cost and predict unstable days better.

Considering all this, including Sk38 and Load seem to offer the preferred classification tree. The trees will also be discussed when including and excluding all three snowpack properties resulting in three considered cases:

Case 1: All variables including Sk38, Load and *H*

Case 2: All variables excluding Sk38, Load and *H*

Case 3: All variables including Sk38 and Load, but not *H*

In the following graphs and figures describing Cases 1, 2 and 3, (a) refers to the learning sample and (b) to the test sample in the classification tree analysis.

The tree in Figure 7.1 (Case 1(a)) is complex, with 18 splits and 19 terminal nodes.

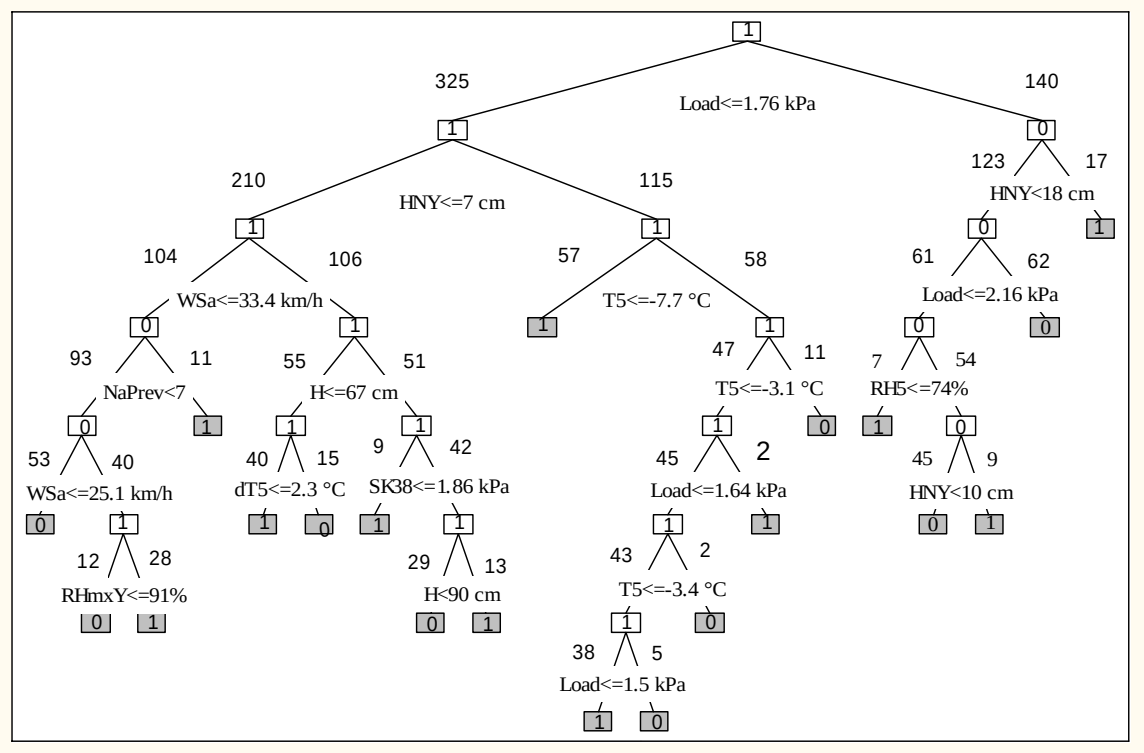

Figure 7.1. Classification Tree Case 1(a) (N = 465). Gray boxes show terminal nodes. The value in the boxes is the predicted DSI class. The number above the box is the number of days to be split.

Load, HnY, RH, T, WSa, NaPrev, RHmax, dT, *H* and Sk38 are used as splitting variables in Figure 7.1. The first split uses Load with a classification of DSI = 1 when Load <= 1.76 kPa. Load is used additionally in three other splits, mostly indicating that fewer unstable days are expected when the Load is higher, which is expected, because the higher the load, the deeper the weak layer is most likely to be, and possibly harder to trigger. This is also seen in the negative correlation (Table 7.4). HnY is used in three splits, always leading to unstable days when the value is higher, which is expected and is discussed in the Spearman rank correlation analysis. Assessing the value of snowpack properties, the split using *H* and Sk38 is interesting. DSI = 1 is classified when *H* > 67 cm and Sk38 < 1.86 kPa but also when Sk38 > 1.86 kPa and *H* > 90 cm, which represents deeper burials in the dataset, considering that on most of the days *H* does not exceed 130 cm and could be that the profile site was different from the conditions at the trigger point.

Similar to the Spearman rank correlations, the predictor importance ranking shows that Load, followed by Sk38, are the most important predictors (Figure 7.2) indicating their influence in a multivariate analysis, but also that *H* is less important in this multivariate analysis.

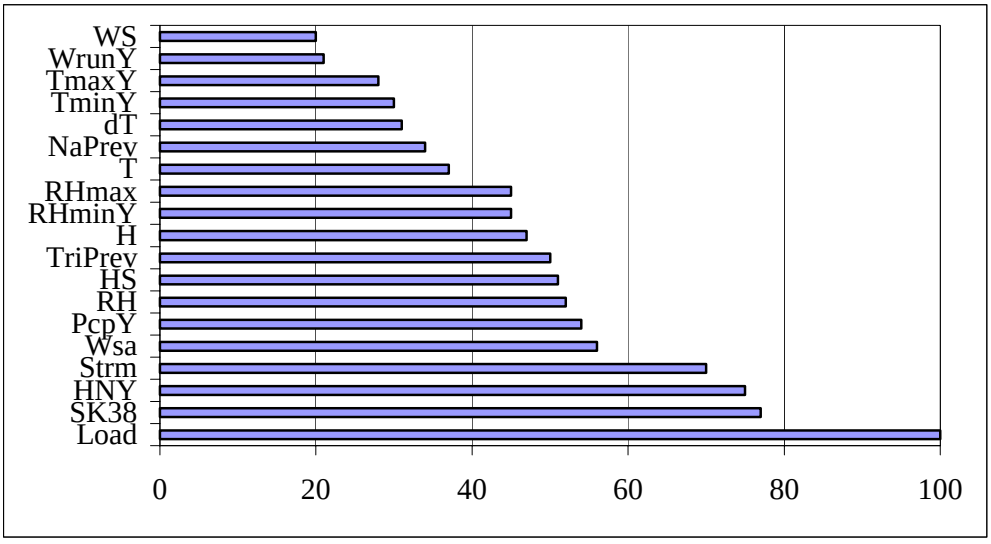

Figure 7.2: Predictor importance ranking including H, Load and Sk38 (Case 1(a)).

According to Figure 7.4 WSa is the most important variable when excluding Sk38, *H* and Load, though the first split in the tree in Figure 7.3 occurs using HnY with more than 7.8 cm of new snow classifying unstable days. Another important splitting variable is TriPrev with four splits, each indicating that the more previously triggered avalanches, the more likely the forecast day is DSI = 1. Though the temperature is not significant in the correlation analysis, three splits in the tree analysis suggest that warmer temperatures lead to unstable days. In addition, RHmax, WS, NaPrev, HS and Strm are used in the tree. Although *T*, NaPrev, WS are not significantly correlated in the univariate analysis, these variables seem to gain importance when considering their interaction with other parameters, whereby the NaPrev are ranked fifth in the importance plot in Figure 7.4.

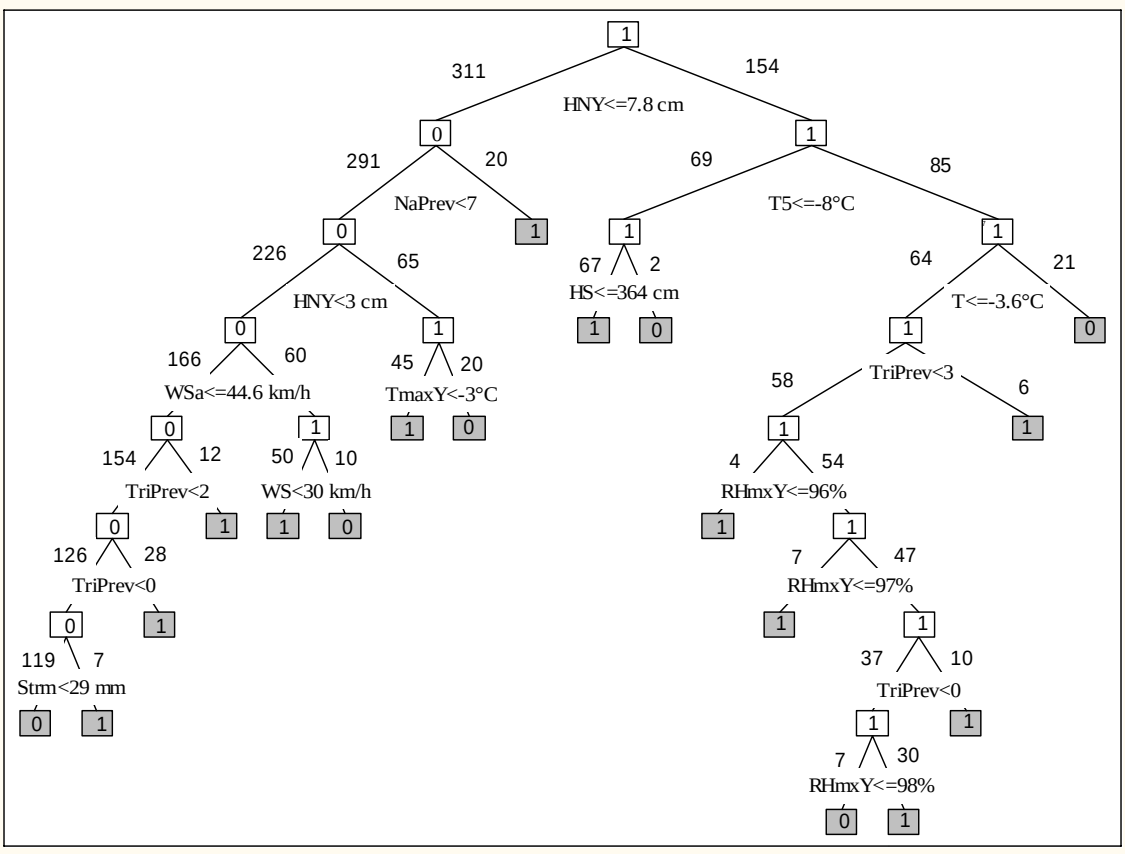

Figure 7.3: Classification Tree Case 2(a) (N = 465). Gray boxes show terminal nodes. The value in the boxes is the predicted DSI class. The number above the box is the number of days to be split.

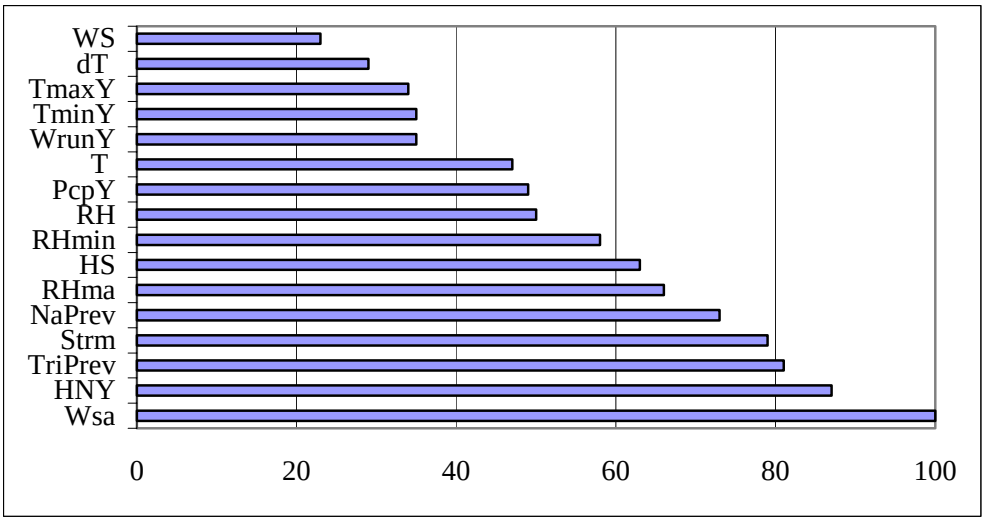

Figure 7.4: Predictor importance including H, Load, SK38

The classification tree with Load and Sk38, but not *H*, has a similar structure to the first tree (including all three predictors). Load is also used in the first split and the tree adds no new solution to the question of whether snowpack properties have value to predict DSI = 1 and consequently is not graphed. As shown before, *H* is not as good a predictor as in the previous study and does not gain importance in the multivariate analysis. In any case the same three datasets will be analysed using the Nearest Neighbour model in the next section, because different techniques may use different variables.

The results from the accuracy and skill measures (see Section 4.3.4) are shown in Figure 7.5 for the datasets including Sk38, Load and *H* (Case 1), excluding Sk38, Load and H (Case 2) and including Sk38 and Load but not *H* (Case 3). (a) refers to the learning sample and (b) to the test sample.

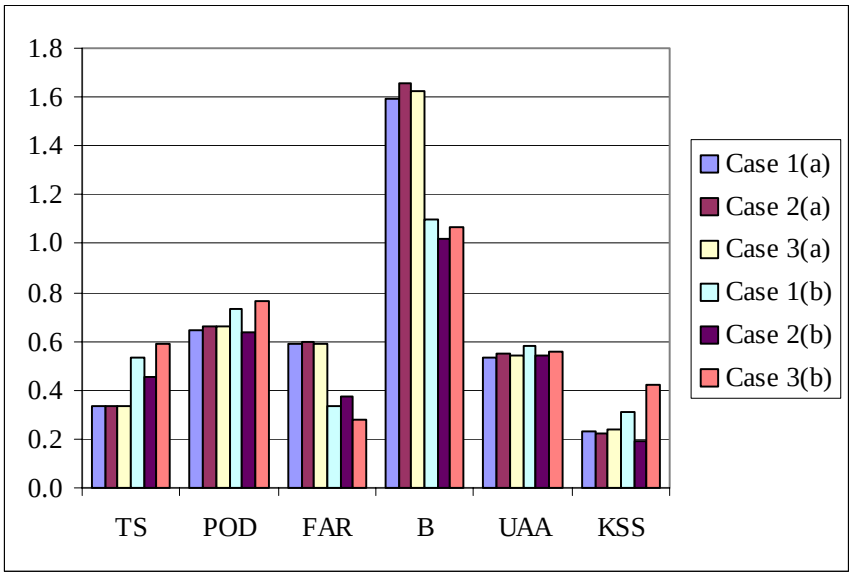

Figure 7.5: Forecast verification. Accuracy and skill measures for the results from the classification tree analysis. Blue River.

The threat scores (TS) in all learning samples were similar with 0.33 in Cases 1 and Case 2 and 0.34 in Case 3 indicating that only about 33% of the unstable days were correctly classified. However the unbalanced dataset with 163 of 481 (34%) unstable days in the learning sample may have influenced the score. The test sample is with 64 of 115 (56%) unstable days more balanced, but also consists of more unstable days than stable days, consequently a better TS is expected in the test sample. The highest score was calculated for Case 3(b) implying that unstable days were most accurately predicted.

The probability of detection (POD) was again similar in all learning samples with 0.65 for Case 1(a) and 0.66 for Cases 2(a) and 3(a). The results in the test sample were better in Case 1(b) with 0.73 and 0.77 in Case 3(b), than in Case 2(b) with 0.63. Considering that the highest possible score is 1 and the higher the score the higher the fraction of correctly forecast unstable days the results in the test sample are surprisingly high compared to the score in the learning samples, possibly a result of the more balanced dataset.

The false alarm (FAR) in all the learning samples is high with 0.59, 0.6 and 0.59 for Case 1, Case 2 and Case 3 respectively. The results in the test samples are again better, but

with in Case 2(b) (0.38) suggesting that the exclusion of the three snowpack properties caused a higher prediction of unstable days on days when no avalanches occurred.

The bias (B) indicates that all Cases overforecast unstable days in the learning sample, again a result of the unbalanced dataset. In the test samples the bias is only slightly above 1 indicating that only a few more unstable days were predicted than occurred. However this does not mean that all unstable days were correctly forecast. Indeed the misclassification of unstable days in the test sample was 27%, 37% and 23% for Cases 1, 2 and three respectively (Table 7.5).

A Kuipers Skill Score (KSS) close to zero indicates a random forecast. The values in the learning sample are 0.23, 0.23 and 0.24 for Cases 1-3. Again the forecast seems not to be satisfactory considering that the best score is one. The best KSS in the test sample was calculated for Case 3(b) with 0.42 and the worst with 0.19 for Case 2(b). As in the other measures excluding Sk38, Load and *H* led to a less accurate forecast than when including these three variables.

Even though the results for Case 2(a) are similar to the results from Case 1(a) and 3(a) the test sample the accuracy measures and the skill scores indicate a less good prediction without Sk38, Load and *H*. The performance of the models in the learning samples is comparable in all three cases. However the poor performance of Case 2(b) indicates that including Sk38, Load and optionally *H* improved the forecast, which confirms the results in Table 7.5.

7.3.3 Nearest Neighbour

A major difference between the Nearest Neighbour model and the Classification Tree analysis is that in the latter, the priors, which can be set in order of importance for forecasting one avalanche class more accurately, cannot be set in the Nearest Neighbour software Cornice. Instead Cornice accounts for the importance of correctly forecasting unstable days and the generally unbalanced dataset (more stable days reported than unstable days) by saying the forecast is correct on a particular day if three or more of the ten most similar days have DSI = 1 and DSI = 1 occurs on the forecast day, or if fewer than three of the most similar days have DSI = 1 and DSI = 0 occurs on the forecast day.

As outlined in Section 4.3.2, a function of Cornice allows the weights of the predictor variables to be set manually or automatically through a generic algorithm. The automatic weights as calculated by Cornice give no physical explanation of the predictors, but have the advantage that they are easy to configure. In Blue River, three senior guides at a heli-skiing operation were asked to weight the predictor variables in regard to their importance in the stability evaluation. The average weights for each predictor variable were used in one dataset to compare the performance of the automatic weighting function and the weights set by expertise and local knowledge.

The configuration process of Cornice was described in Section 4.3.2. Four different datasets were inputted separately:

Case 1:	All variables including Sk38, Load and *H*
Case 2:	All variables excluding Sk38, Load and *H*
Case 3:	All variables including Sk38 and Load, but not *H*
Case 4:	All variables including Sk38, Load and *H* using the weights set by guides

In the following graph and figures describing Cases 1, 2, 3, and 4, (a) refers to the learning sample, (b) to the test sample and (c) to the total dataset. Cases (b) and (c) are configured using the scales and weights from the learning sample.

Due to missing input data, 51 of 596 days were not forecast. Table 7.6 summarizes the results from each configuration for the learning sample, the test sample and the total dataset using the configuration of the learning sample (scales and weights).

Table 7.6. Performance of Nearest Neighbour model. Blue River.

			Unstable days (DSI = 1)			Stable days (DSI = 0)			
		Days forecast	Correct	Wrong	% correct	Correct	Wrong	% correct	Fitness of batch test (%)
Case 1	(a)	439	96	30	76	218	95	70	73
	(b)	106	44	11	80	39	12	77	78
	(c)	545	139	42	77	245	119	67	72
Case 2	(a)	439	88	38	70	192	121	61	66
	(b)	106	43	12	78	25	26	49	64
	(c)	545	136	45	75	198	166	54	65
Case 3	(a)	439	94	32	75	210	103	67	71
	(b)	106	43	12	78	23	28	45	62
	(c)	545	135	46	75	240	124	66	70
Case 4	(a)	439	73	53	58	210	103	67	63
	(b)	106	43	12	78	44	7	86	82
	(c)	545	126	55	70	254	110	70	70

The best result in the learning sample was achieved using Sk38, Load and *H* as predictor variables, with a fit of 73%, and the worst result occurred using the weights of the predictors set by the guides. These results show that Sk38, Load and *H* improved the forecast in the learning sample by 7% compared to using none of the three predictors, a much larger difference than in the tree analysis where Sk38, Load and *H* improved the forecast in the learning sample by only 1% based on the number of misclassified days.

In Case 3, *H* was excluded from the analysis because a better classification of unstable days was achieved without *H* in the classification tree analysis. Also, the Spearman rank correlation analysis showed only a weak correlation of DSI with *H*. Other than in the classification tree, the Nearest Neighbour forecast in these three samples led to a less accurate fit in both stable and unstable forecasts compared to Case 1. Thus it can be concluded that *H* does have a predictive value in a Nearest Neighbour approach. As a result, Case 3 is excluded from further analysis.

The test sample was best predicted in Case 4 with 82% fit, though the learning sample yielded the least accurate fit of all four cases. Purves et al. (2003) mentioned that when forecasters determine the weights they tend to set them according to the snow and avalanche conditions during the more recent winters. This could have led to the good fit of the test sample, because it consists of data from the past two winters. The second best fit of

the test sample was observed in Case 1 with 78% fit. The improvement compared to Case 2 was 14% and again Sk38, Load and *H* improved the forecast substantially.

In the total dataset, which is a combination of the learning sample and the test sample, the fit was 1% less accurate than in the learning sample on its own, except for Case 4, indicating that the additional years in the datasets did not improve the forecast of the years in the learning sample. It seems that the past two winters (test sample) were somewhat different from the winters used in the dataset.

In Section 4.3.4 accuracy and skill measures to verify forecasts are introduced. Figure 7.6 gives an overview of the results for the Cases 1, 2 and 4.

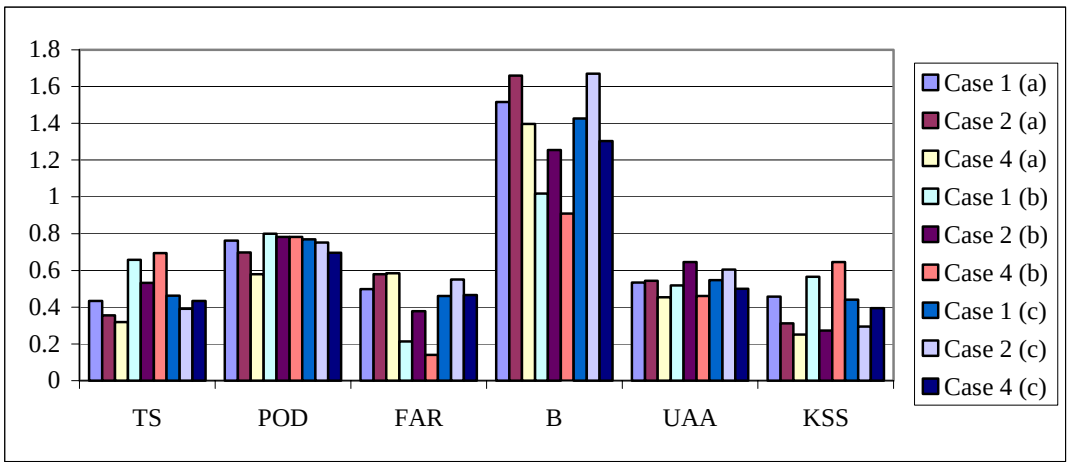

Figure 7.6: Forecast verification. Accuracy and skill measures from the results form the nearest neighbour analysis. Blue River.

In the learning sample the best threat score (TS) was calculated for Case 1 with a value of 0.43. However, considering that the best possible TS is 1, indicating that all avalanche days would have been correctly forecast, the score is not particularly high. In the test samples the scores yielded more promising values with 0.69 (Case 4 (b)), 0.66 (Case 1 (b)) and 0.53 (Case 2 (b)). However, the improved scores are influenced by the balance of the datasets since the test sample has a ratio of 1.1:1 unstable to stable days compared to a ratio of 0.4:1 in the learning sample.

The probability of detection (POD) is the fraction of correctly forecast avalanche days with a perfect forecast being 1. Because in avalanche forecasting it is important to reduce false stable predictions, this score is important for verification and thus high values are desirable. In all cases but Case 4 (a), the scores are 0.7 or higher indicating that DSI = 1 is predicted quite accurately. Again the best scores were computed for Case 1.

The false alarm (FAR) is a measure of false unstable predictions. Even though the consequences of a high misclassification in this forecasting score are lower than the consequences of misclassification calculated in the POD, an unreasonably high score would provoke mistrust in the forecasting model. A score of zero means that no false unstable predictions were made. The score in Case 1 is lowest with a value of 0.5 indicating that stable days were predicted wrongly on half of the days. In Cases 2 and 4 the prediction of false unstable days is even higher. However, this score is highly biased because of the unbalanced datasets in the learning sample and total dataset. In the test sample the values were better with 0.21, 0.38 and 0.14 for Cases 1, 2 and 4 respectively.

The bias (B) is a measure of the average forecast, comparing the number of forecast days with the number of observed days. A bias greater than 1 indicates that the number of forecast unstable days is higher than the number of unstable days in the observations. In most cases more unstable days were forecast than observed, except in Case 4 (b). This is a reasonable result because it indicates that more false unstable days are likely to be predicted than false stable days. In the test sample the bias is close to 1 for Cases 1 and 4, indicating that the numbers of predicted and observed unstable days are almost equal. The bias has to be interpreted carefully because it is an average measure and even though the numbers of predicted and observed unstable days may be equal this does not mean that all the unstable days were correctly forecast.

The Kuipers Skill Score (KSS) compares the set of forecasts to a random forecast and is best when the score equals 1. Zero indicates that the forecast is random and a score of less than zero is an even a less accurate forecast. Case 4 (b) yielded the best score of 0.65 and Case 4 (a) yielded the lowest score of 0.25. Again the best scores were calculated in the test sample with a balanced dataset.

These skill and accuracy scores have to be interpreted with the absolute results in Table 7.6. A decision on whether the forecast quality is good or bad cannot be made based on these measures alone. However, it can be concluded that the quality of Case 1 is the best in most respects. This result is a little different from the classification tree analysis where the forecast in the learning sample was comparable for Case 1(a) and 2(a). However as in the classification tree the test sample for Case 1 was better predicted than in Case 2. Comparing the overall performance of the classification tree analysis and the Nearest Neighbour model the accuracy measure and the skill score indicate a better forecast using the Nearest Neighbour model with higher TS, POD and KSS values and lower FAR. However the bias is almost comparable.

Case 1 yielded the best fit in the Nearest Neighbour analysis for the learning sample and the total dataset and the second best fit in the test sample. Because of the poor fit of the model in Case 4 with the learning sample, Case 1 is thought to be the most promising configuration for the Blue River dataset and will be analysed more closely. However, it is interesting to compare the misclassified days of the dataset both using all three snowpack and stability predictors and then without these predictors, to analyse the value the predictors added to the forecast.

Next, the forecast of three different weak layers, which seemed to represent typical time series, are discussed. Figure 7.7a (including Sk38, Load and *H*) and 7.7b (excluding Sk38, Load and *H*) represent the forecast of Cornice for the faceted layer buried on the 22nd of November 1996, using the weights from the learning sample. The columns above the x-axis represent DSI = 0 observations and below, DSI = 1 observations. Green columns represent correct forecasts and red columns are wrong forecasts. Consequently, red columns above the x-axis represent false unstable conditions – an avalanche day was forecast, but no avalanches occurred - and the red columns below the x-axis are false stable predictions. Sk38, Load and *H* improved the forecast significantly, with one false unstable and two false stable predictions compared to 13 false unstable and two false stable predictions. However, both configurations did not predict some of the avalanches observed in the second half of the observation period.

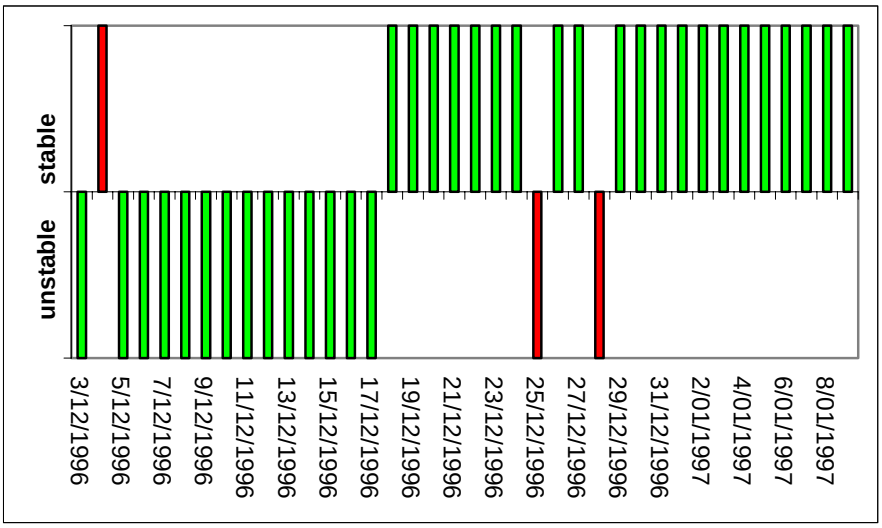

Figure 7.7a: 22nd of November 1996 Facets. Case 1.

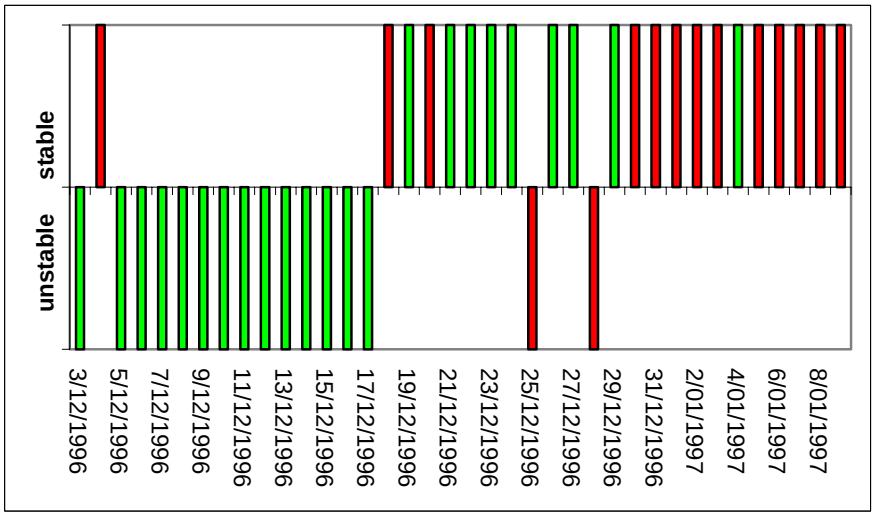

Figure 7.7b: 22nd of November 1996 Facets. Case 2.

This pattern can be observed in other time series as well. Without Sk38, Load and *H* the forecast depends highly on the 24 hour precipitation, humidity and wind speed, all predictor variables important for forecasting storm snow avalanches. Consequently, an unstable day is often forecast when no avalanches occurred on a persistent weak layer, but storm snow avalanches were likely. Including Sk38, Load and *H* reduces this error and leads to a more accurate forecast.

Figure 7.8a and b is the forecast made for the 31st of January 2000 surface hoar layer. The forecasts in Cornice were only possible starting on the 12th of February - almost two weeks after burial - because the calculated ski penetration was deeper than the weak layer until then. Nonetheless, most of avalanche activity was observed after the 12th of February. Including Sk38, Load and *H* misclassified two days fewer than without these predictors and additionally reduced the false stable predictions from four to one.

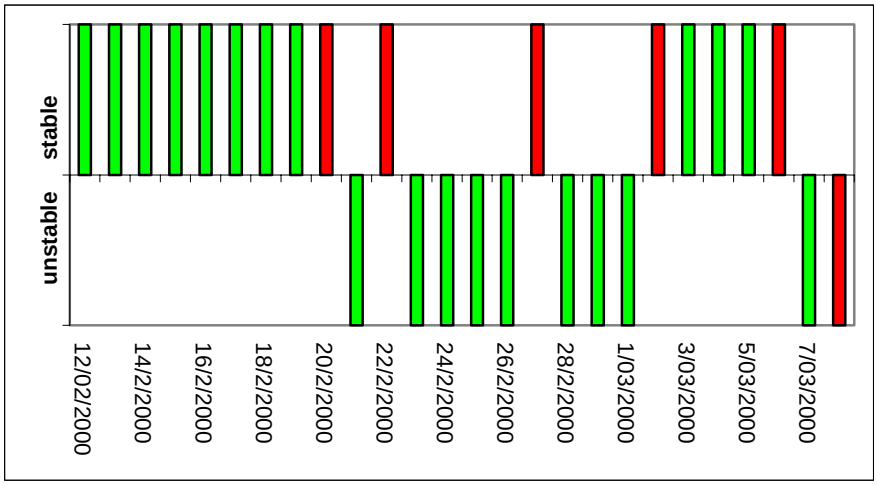

Figure 7.8a: 31st January 2000 surface hoar. Case 1.

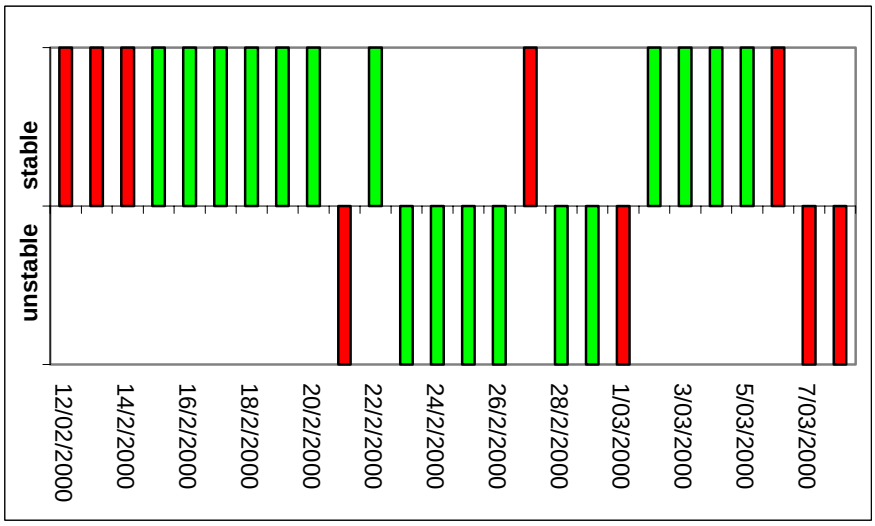

Figure 7.8b: 31st January 2000 surface hoar. Case 2.

The last example shows a faceted layer buried on the 4th of March 2004 where more than 50% of the days were misclassified in both configurations (Figure 7.9a and b). The model failed because of snowpack conditions not represented by the predictor variables. After the faceted layer was buried, a major melt-freeze crust formed on the snow surface and prevented avalanche releases on many slopes. Therefore only a few avalanches were reported after the 15th of March. However, the predictor variable used in Cornice to find the Nearest Neighbour days did not include any information about surface crusts. Consequently, false unstable conditions were forecast. It is expected that this was not the only time series affected by a crust above the weak layer. This is a shortcoming of predictors used in the Nearest Neighbour model.

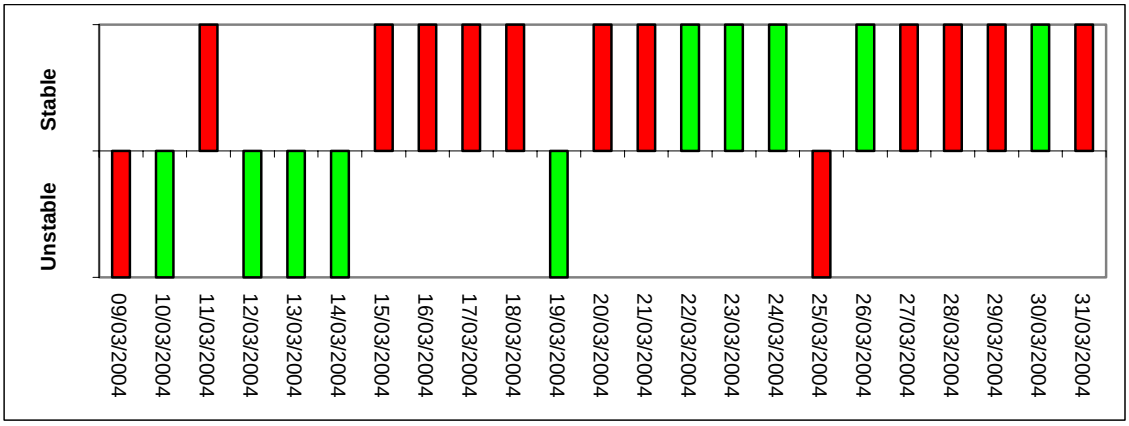

Figure 7.9a: March 4th 2004 surface hoar. Case 1.

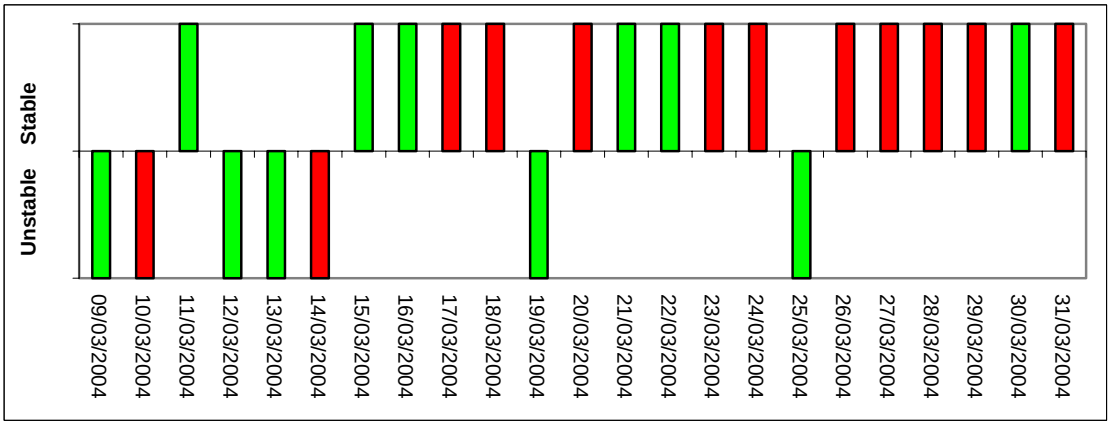

Figure 7.9b: March 4th 2004 surface hoar. Case 1.

All in all the forecast made by Cornice has its values but also its shortcomings. However, the information the forecaster can get from the model seems to be valuable, though the model has to be used as a tool and the information has to be applied carefully to the daily stability information. Forecasters and guides most value the list of avalanches that occurred on similar days.

7.3.4 Incorporation of rutschblock scores

Rutschblock tests are commonly done at an intermittent basis throughout the study area. The idea is to estimate the shear strength from Rutschblock scores and calculate Sk38. If the calculated Sk38 is lower than the one calculated using shear strength extrapolations based on shear frame measurement then the lowest value is chosen in order to improve the forecast of unstable days.

Jamieson (1995: 177) showed that Sk (Equation 2.9) can be estimated from rutschblock scores and found a relation of

$$Sk = 0.31(RB-1) \qquad \textit{Equation 7.3}$$

Sk is a function of the shear strength of a particular weak layer, which can be estimated using Equation 2.9. The calculated shear strength was adjusted to normal load and then Sk38 calculated. It was common practice of ASARC to record the ski penetration after two jumps. Consequently in Equation 2.8, Pk was taken as the recorded ski penetration after two jumps instead of calculating the penetration based on the slab density. Out of a total of 70 rutschblocks, 36 yielded lower values of Sk38 than the value for the same day extrapolated from study plot measurements.

The results of the nearest neighbour model after updating Sk38 are shown in Table 7.7.

Table 7.7: Performance of Nearest Neighbour model using RB scores. Blue River.

		Days forecast	Unstable days (DSI = 1) Correct	Wrong	% correct	Stable days (DSI = 0) Correct	Wrong	% correct	Fitness of batch test (%)
SkRB	(a)	439	90	36	71	219	94	71	71
	(b)	106	41	14	75	33	18	65	70
	(c)	545	127	54	70	232	132	64	67

The overall forecast including days where Sk38 was estimated from rutschblock scores compared to the results from the shear strength calculations in Section 7.3.3 is less accurate with 71%, 70% and 67% for (a), (b) and (c) respectively compared to 73%, 78% and 72% in Table 7.6 Case 1 (a), (b) and (c). Also the forecast for unstable days did not improve. In the learning sample six more false stable days were forecast, in the test sample 3 and in the combined 12 when including days when Sk38 was estimated from the rutschblock scores.

The forecast verification in Figure 7.10 confirms the results. The scores are graphed for the cases using the rutschblock scores (RB (a)-(c)) along with the results not using RB ((a)-(c)) in Section 7.3.3 corresponding to Case 1 (a)-(c).

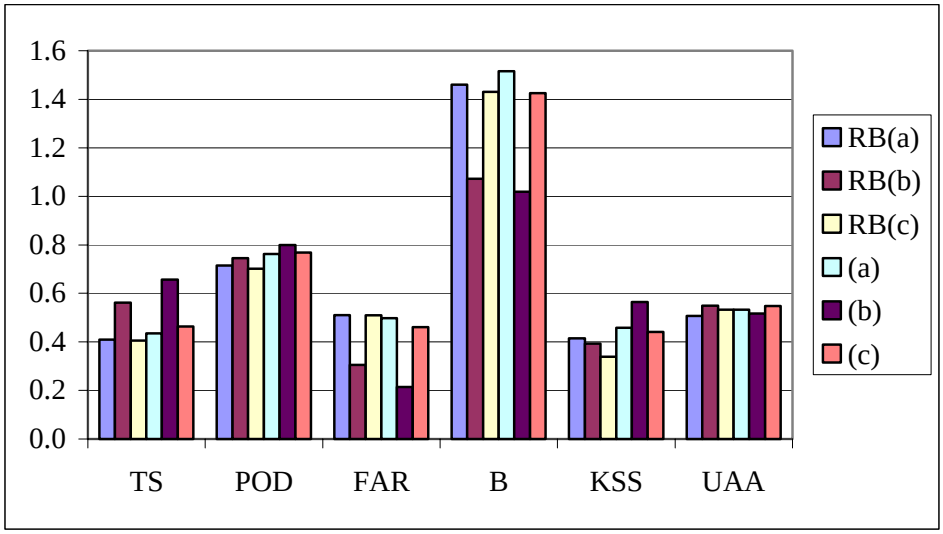

Figure 7.10: Forecast verification of nearest neighbour analysis. Comparison with and without using rutschblock scores to estimate the shear strength. Blue River.

In all three samples, the threat score (TS) and the probability of detection (POD) indicate that the inclusion of the RB scores provoked a less accurate forecast for unstable days. The same was observed for the unweighted average though the test sample in RB(b) yielded a slightly better result than Case 1 (b). Additionally the false alarm rate was higher when including the rutschblock score results showing that the weaker values for Sk38 did not reflect the observed conditions on the regional scale. RB(a) and (c) and Case 1(a) and (c) are overforecast whereas the two test sample are close to a bias (B) of one indicating that the same amount of unstable days were forecast and occurred. Although this does not mean that all of the unstable days were correctly forecast. Case 1(b) yielded a fit 80% and RB(b) a fit of 75%. The Kuipers skill score (KSS) indicates a better forecast without using the RB results also showing that the inclusion did not add value to the prediction.

The forecast using the rutschblock score to calculate the shear strength and Sk38 did not improve the regional forecast in this example. However, rutschblock tests on slopes or in areas that lack data may still be useful in forecasting operations. This sample dataset is only the first approach to assessing the value of rutschblock tests at various locations away from the study plot in a regional forecasting model. Nevertheless, the results in this section, combined with the known variability of rutschblock scores in avalanche terrain (Campbell, 2004), question the merit of applying rutschblock scores from avalanche terrain to regional forecasts. Jamieson (1995: 180-182) showed that rutschblock scores from study slopes correlated with the frequency of skier triggering within 15 km. Again, this shows the predictive value of stability indices such as Sk38 and RB for regional avalanche forecasting, provided the indices are obtained in representative study slopes or plots.

In all three samples, the threat score (TS) and the probability of detection (POD) indicate that the inclusion of the RB scores provoked a less accurate forecast for unstable days. The same was observed for the unweighted average though the test sample in RB(b) yielded a slightly better result than Case 1 (b). Additionally the false alarm rate was higher when including the rutschblock score results showing that the weaker values for Sk38 did not reflect the observed conditions on the regional scale. RB(a) and (c) and Case 1(a) and (c) are overforecast whereas the two test sample are close to a bias (B) of one indicating that the same amount of unstable days were forecast and occurred. Although this does not mean that all of the unstable days were correctly forecast. Case 1(b) yielded a fit 80% and RB(b) a fit of 75%. The Kuipers skill score (KSS) indicates a better forecast without using the RB results also showing that the inclusion did not add value to the prediction.

The forecast using the rutschblock score to calculate the shear strength and Sk38 did not improve the regional forecast in this example. However, rutschblock tests on slopes or in areas that lack data may still be useful in forecasting operations. This sample dataset is only the first approach to assessing the value of rutschblock tests at various locations away from the study plot in a regional forecasting model. Nevertheless, the results in this section, combined with the known variability of rutschblock scores in avalanche terrain (Campbell, 2004), question the merit of applying rutschblock scores from avalanche terrain to regional forecasts. Jamieson (1995: 180-182) showed that rutschblock scores from study slopes correlated with the frequency of skier triggering within 15 km. Again, this shows the predictive value of stability indices such as Sk38 and RB for regional avalanche forecasting, provided the indices are obtained in representative study slopes or plots.

7.4 Glacier National Park and Adamants

The second dataset is a combination of meteorological and snowpack information measured in Glacier National Park and the Adamants helicopter skiing area (Figure 3.1, Chapter 3). Meteorological data were available from Mt. Fidelity in Glacier National Park at a commonly skied altitude of 1905 m about 50 km south of the Adamants. Additionally shear frame measurements were made about 150 m from the weather station on a northeast exposed study slope with an inclination of about 25-38°. Additionally, CMH provided

meteorological data from a study plot at 1100 m elevation in close vicinity to the ski runs. CMH Adamants also provided the avalanche occurrence data.

The entire dataset consisted of observations from the winters 1999/2000 – 2003/2004 and is much smaller than the Blue River dataset with only 321 days measured in 10 time series. CMH Adamants usually do not start their season before January and consequently observations on the November facet layers are rare. Most avalanche observations are on persistent weak layers, which formed between the end of December and the end of March. Shear frame measurements on persistent weak layers are generally not available in April because University of Calgary snowpack observations on Mt. Fidelity do not continue after the end of March. Even though some facet layers were tested during the winter these were never the weakest in the study plot and consequently the entire dataset consists of measurements on surface hoar layers. Three time series were excluded for use as a test sample and the remaining seven were used as the learning sample (Table 7.8).

Table 7.8: Surface hoar time series measurements at Mt. Fidelity. Bolded italics indicate the test sample.

WL date	N
31-Dec-1999	33
21-Feb-2000	33
28-Jan-2001	30
23-Feb-2001	29
2-Jan-2002	34
16-Feb-2002	36
25-Dec-2002	30
20-Jan-2003	***26***
15-Feb-2003	***32***
14-Feb-2004	***38***

Out of a total of 321 days, 83 were assigned DSI = 1 where an actual avalanche occurred on 66 days and 17 days were rated as days with poor stability but no actual avalanches were reported on a persistent weak layer. The remaining 238 days where DSI = 0.

The availability of meteorological and snowpack variables suggests the analysis of four different predictor variable combinations with DSI:

Case 1:	Meteorological data and snowpack properties from Mt. Fidelity.
Case 2:	Only meteorological data from Mt. Fidelity.
Case 3:	Meteorological data from the Adamants and snowpack properties from Mt. Fidelity.
Case 4:	Only meteorological data from the Adamants.

The results for the learning and tests samples are labelled (a) and (b), respectively. For example, the results for Case 1 are labelled Case 1(a) and 1(b).

By using these different sets of predictor variables it is possible to assess the predictive value of the snowpack properties H, Load and Sk38 and also to analyse if these variables can improve the forecast in an area about 50 km north from the measurement site.

The two meteorological datasets are quite different in nature. At Mt. Fidelity the measurements are recorded by an automatic weather station, whereas the data in the Adamants are from manual weather observations. Wind speed, the sky value, precipitation and the change of barometric pressure are observed qualitatively and expressed as number (Table 7.1). These variables can be used in the analysis since nonparametric methods were chosen which do allow that the level of measurement of the variable being nominal or ordinal (Burt and Barber, 1996: 331). In the classification tree analysis a distinction can be made into categorical and ordered predictors.

7.4.1 Spearman rank correlations

The results of the Spearman rank correlations can be seen in Table 7.9 for the Mt. Fidelity dataset and in Table 7.10 for the Adamants dataset. Both tables include Sk38, Load and H to compare their ranking in relation to the other predictor variables. In both datasets Sk38, Load, TriPrev and H correlate highest with the response variable DSI, indicating their importance to forecast skier-triggered avalanches on persistent weak layers. However a further comparison is difficult, because the predictors are different. In both datasets HN

was not significant, which is surprising because in the Blue River dataset the new snow height had the fourth strongest correlation with DSI. In the dataset for the Adamants this is explicable because the elevation of the study plot is 1100 m, well below common skiing elevations, and some snowfall at skiing elevations may fall as rain at the study plot in March. Comparing the correlation from Mt. Fidelity with the results from Blue River (Table 7.4) it is notable that most of the correlations are similar, indicating that the weather and snowpack conditions are similar over this part of the Columbia Mountains. The explanations for the correlations with DSI can be reviewed in Section 7.3.1. Temperature is generally considered an important predictor for avalanching, however none of the temperature variables were significant in the Mt. Fidelity and the Blue River datasets, possibly because storm snow avalanches were excluded. Interesting is that temperature was more strongly correlated in the Adamants dataset.

Table 7.9. Spearman rank correlations with DSI. Insignificant variables (p > 0.05) are in italics. Mt. Fidelity.

Variable	Valid N	Spearman R	p-level
Sk38	321	-0.413	1.2E-14
Load	321	-0.390	4.3E-13
TriPrev	321	0.241	1.3E-5
H	321	-0.238	1.7E-5
Strm	299	0.188	0.001
HS	321	-0.170	0.002
RHmnY	321	0.166	0.003
PcpY	320	0.151	0.007
RHmxY	321	0.129	0.021
WV	320	0.127	0.024
RH	320	0.124	0.027
WS	320	0.117	0.036
HNY	*321*	*0.096*	*0.086*
WrunY	*321*	*0.046*	*0.415*
TminY	*321*	*0.043*	*0.440*
WD	*320*	*-0.650*	*0.516*
T5	*320*	*0.025*	*0.655*
TmaxY	*321*	*0.015*	*0.789*

Table 7.10. Spearman rank correlations with DSI. Insignificant variables (p > 0.05) are in italics. Adamants.

Variable	Valid N	Spearman R	p-level
Sk38	321	-0.413	1.2E-14
Load	321	-0.390	4.3E-13
TriPrev	321	0.241	1.3E-5
H	321	-0.238	1.7E-5
HSTM	304	0.196	6.0E-4
T	304	0.144	0.012
Tmin	304	0.140	0.015
Baro	304	-0.139	0.016
HSTD	304	0.118	0.041
PcpTi	*304*	*0.111*	*0.054*
HN24	*304*	*0.110*	*0.056*
HS	*304*	*-0.101*	*0.079*
Sky	*304*	*0.095*	*0.097*
Tmax	*304*	*0.071*	*0.220*
WS	*304*	*-0.046*	*0.427*
Btrend	*304*	*0.019*	*0.742*

7.4.2 Classification tree analysis

Table 7.11 summarizes the global cross-validation results for the Cases 1-4 for both the learning and the test sample.

Table 7.11. Global cross-validation results. Classification tree. Priors 0.45/0.55 (DSI 0/1).

		Unstable days (DSI = 1)			*Stable days (DSI = 0)*				
	N	*Correct*	*Wrong*	*% correct*	*Correct*	*Wrong*	*% correct*	*Cost*	*Error*
Case 1 (a)	221	28	18	61	139	28	79	0.308	0.042
(b)	78	15	16	48	38	9	81	0.370	0.056
Case 2 (a)	221	27	19	59	95	80	54	0.433	0.043
(b)	78	14	17	45	30	17	64	0.464	0.058
Case 3 (a)	225	30	16	65	147	32	82	0.272	0.041
(b)	78	27	10	73	30	11	73	0.267	0.050
Case 4 (a)	225	18	28	39	149	30	83	0.410	0.042
(b)	78	14	23	38	37	4	90	0.385	0.048

Case 1 is compared to Case 2, and Case 3 to Case 4 to assess whether Sk38, Load and H improved the forecast. In both comparisons the misclassification cost was reduced

significantly by incorporating the three snowpack properties. The cost in the learning sample was reduced from 0.433 to 0.308 and in the test sample from 0.464 to 0.370 in the test sample (Case 1 and 2). However the forecast of unstable days only improved by one correctly forecast unstable day in both samples. The reduction in the cost was mainly influenced by a better classification of stable days. The opposite was observed using meteorological data from the Adamants (Cases 3 and 4) where the misclassification cost was also reduced significantly but as a result of better forecasts for unstable days. Without the three snowpack properties, unstable days were only correctly predicted in 39% of the cases whereas after the incorporation 65% were correctly forecast, though the percentage of correctly predicted stable days was reduced from 83% to 82%. In any event it has been shown in both comparisons that the forecast was improved significantly and that Sk38, Load and H have great predictive potential. Using Sk38, Load and H on their own or in any other combination did not improve the forecast further, which is similar to the results from Blue River (Section 7.3.2).

The lowest misclassification cost was achieved when using the meteorological data from the Adamants in combination with the snowpack observations from Mt. Fidelity (Case 3). However the cross-validations do not reveal the structure of the classification. Comparing Classification Trees (Case 1(a)) (Figure 7.11) and (Case 3(a)) (Figure 7.12) reveals a similar structure.

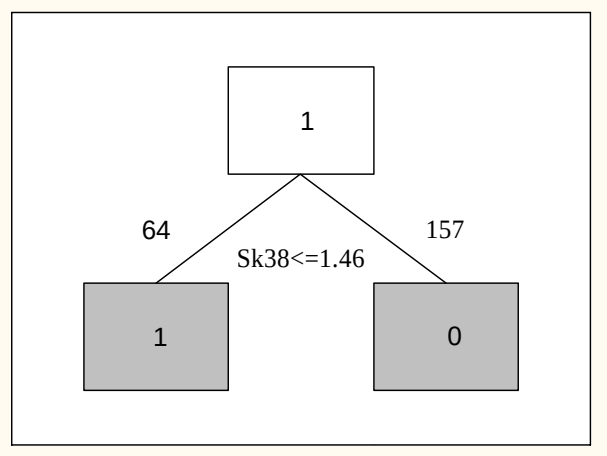

Figure 7.11: Classification tree including Sk38, H and Load. Mt Fidelity (Case 1(a)).

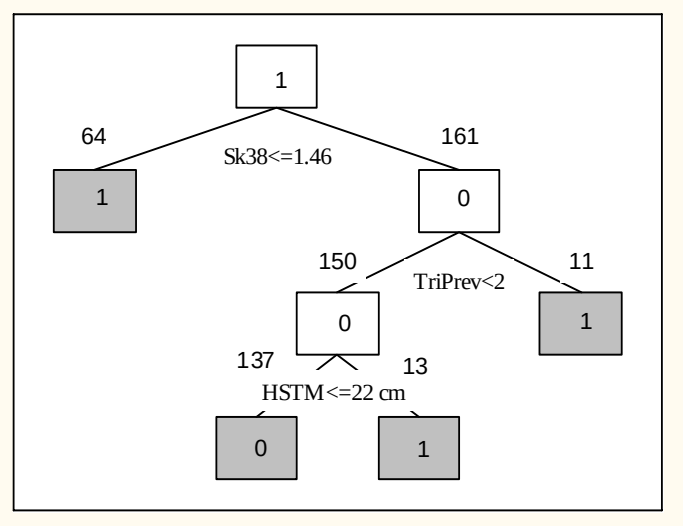

Figure 7.12: Classification tree including Sk38, H and Load. Adamants (Case 3(a)).

In both trees the first split is determined by Sk38, predicting unstable days when the stability index is smaller than 1.46 (left branch). The right branch of the Classification Tree (Case 3a, Figure 7.12) was again split and unstable days were predicted when more than two avalanches on the previous day were skier-triggered. Even though TriPrev was a predictor variable in Case 1(a) this split is absent in Figure 7.11, possibly because four days were not considered because of missing values in the predictor variables in Case 1(a). One last split in Figure 7.12 predicts unstable days when HSTM was greater than 22 cm (right branch, split 3).

In the importance ranking for both Case 1(a) and Case 3(a) Sk38, Load and TriPrev were ranked highest and H was ranked fourth and fifth in the two datasets respectively (Figures 7.13 and 7.14). H does show some value to predict DSI, which is different from the Blue River dataset where H lost in significance because layers shallower than 30 cm were excluded. A possible explanation might be that this dataset is much smaller and fewer cases were excluded because of their shallowness.

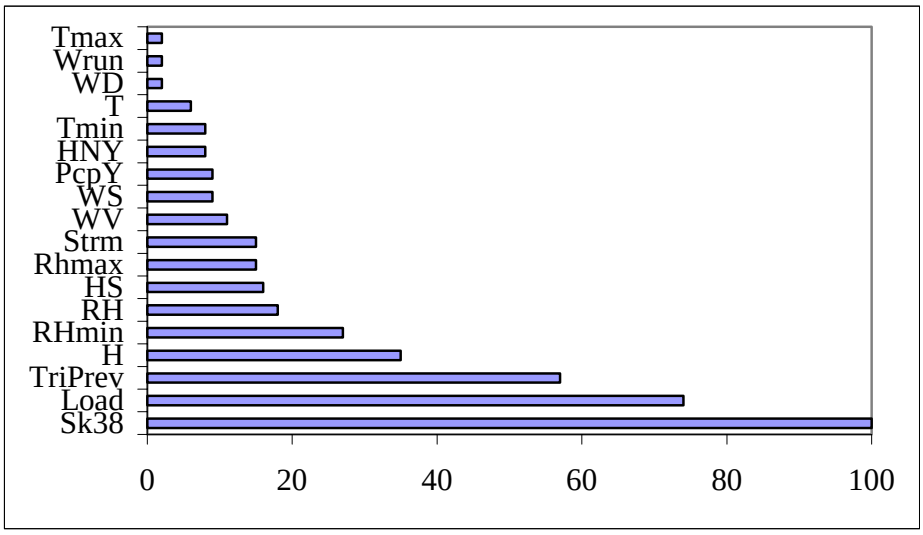

Figure 7.13: Importance ranking. Mt. Fidelity including Sk38, H and Load (Case 1(a)).

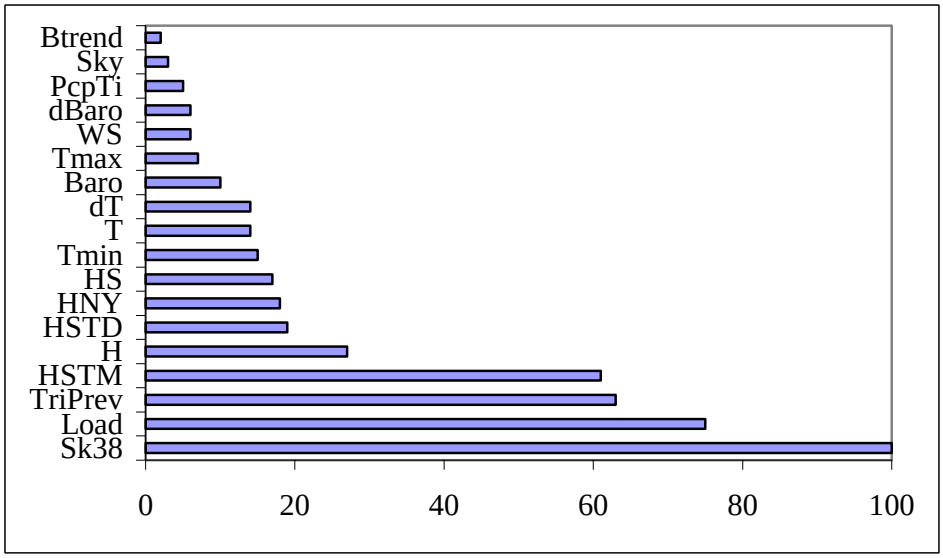

Figure 7.14: Importance ranking. Adamants including Sk38, Load, H (Case 3(a)).

Excluding the snowpack properties H, Load and Sk38 in both datasets (Cases 2(a) and 4(a)) the first split was identical (Figure 7.15 and 7.16).

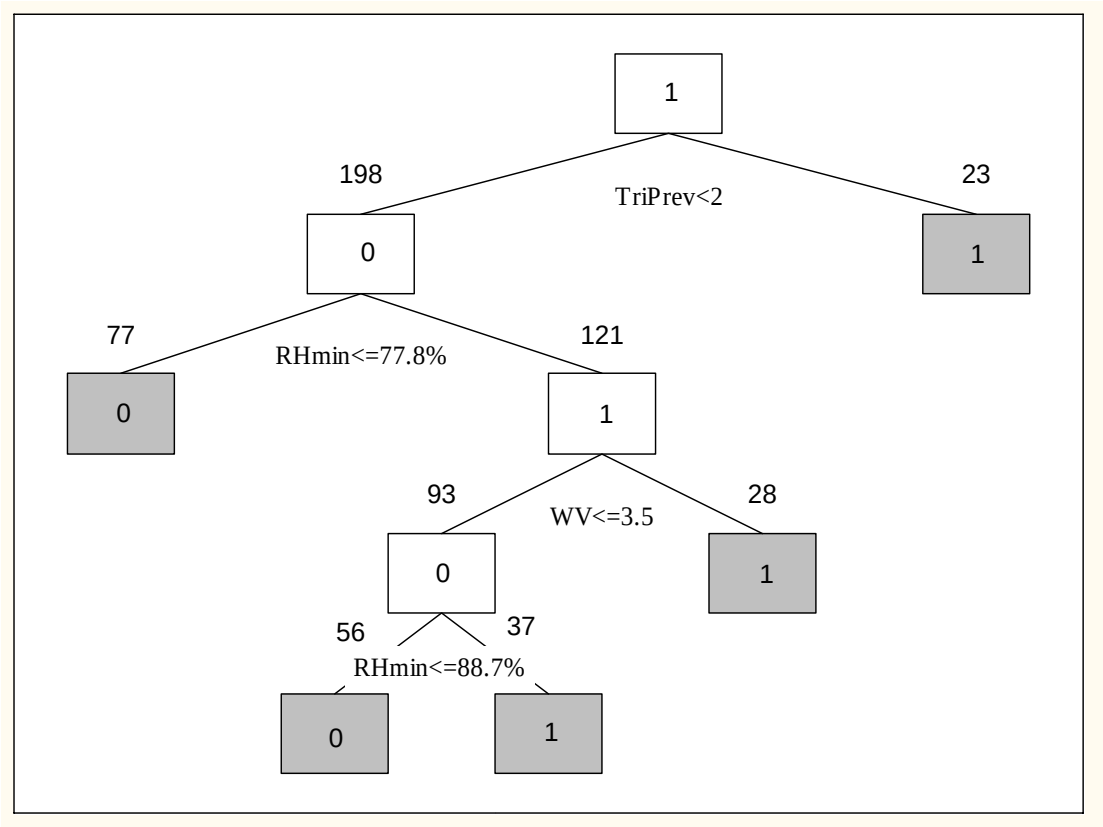

Figure 7.15: Classification tree with Fidelity meteorological data. Excluding Sk38, Load, H. Case 2(a).

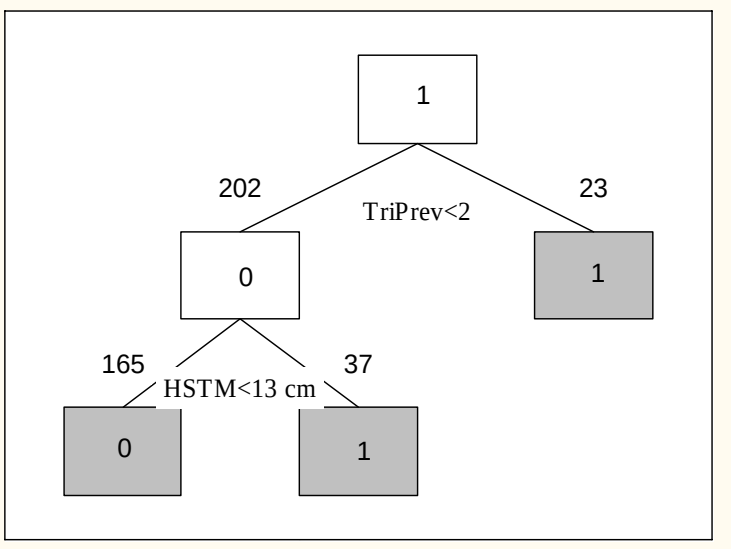

Figure 7.16: Classification tree with Adamants meteorological data. Excluding Sk38, Load, H. Case 4(a).

In the first splits in Figure 7.15 and 7.16 unstable days were predicted when more than two avalanches were skier triggered on the previous day. This is the only variable identical in Cases 2(a) and 4(a). Again Case 4(a) consisted of four more observations than Case 2(a) due to missing values in the meteorological data from Mt. Fidelity. The classification tree in Figure 7.15 shows that with a minimum humidity of less than 77.8% over the past 24 hours stable days were classified (left branch, split two). The right branch of the second split led to another split were unstable days were predicted when the measured wind vector was greater than 3.5 (right branch, split 3). The cases in the left branch were again split by the minimum humidity with higher values than 88.7% classifying unstable days. In Figure 7.16 (Case 4(a)) only one more split with HSTM greater than 13.4 cm unstable days were predicted. The cross-validations in Table 7.11 showed that 39% of the unstable days were correctly classified in Case 4(a) and 59% in Case 2(a) in the learning sample and 38% and 45% respectively in the test sample. This shows that the forecast improved for unstable days using meteorological data from Mt. Fidelity instead of using the data from the Adamants. However stable days were better predicted with 83% compared to 54% in the learning and 90% compared to 64% in the test sample using the meteorological data from the Adamants compared to the data from Mt. Fidelity.

The importance ranking in Figures 7.17 and 7.18 are not as strongly dominated by one variable as was observed in Figures 7.13 and 7.14. Even though TriPrev was used as first splitting variables in the classifications trees in Figures 7.15 and 7.16 this variable is only ranked in fourth position in both plots. The variables used in the second split, HSTM and RHmin respectively were ranked highest.

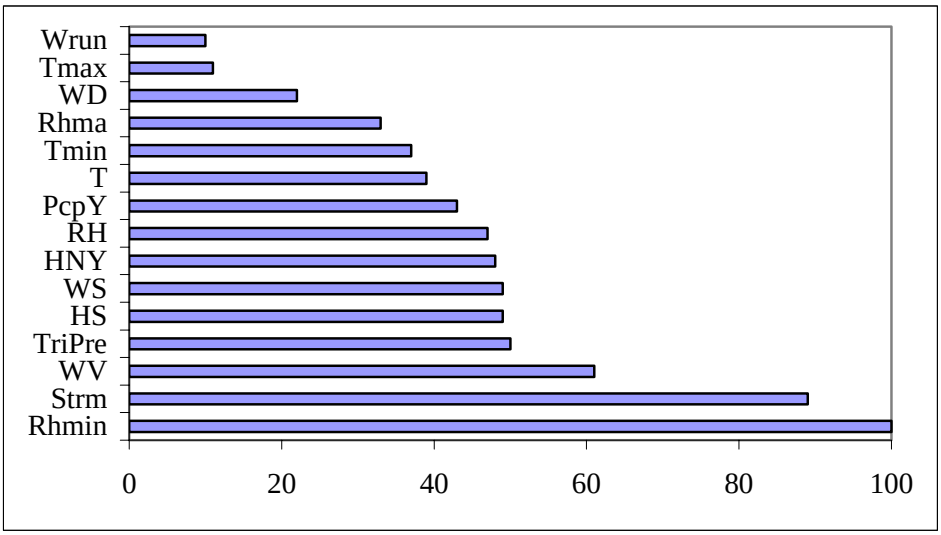

Figure 7.17: Importance ranking with Fidelity meteorological data. Excluding Sk38, Load, H. Case 2(a).

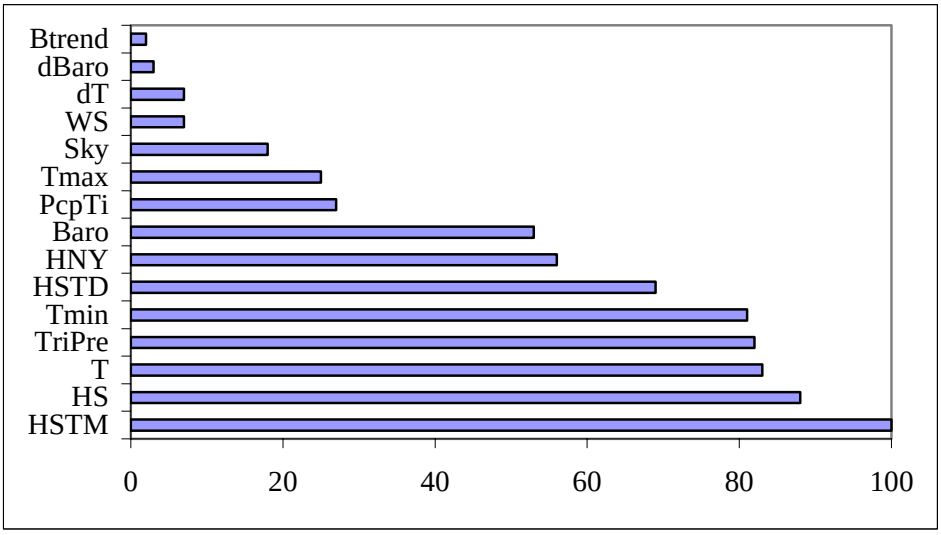

Figure 7.18: Importance ranking with Adamants meteorological data. Excluding Sk38, Load, H. Case 4(a).

The results of the forecast verification for each of the four cases are summarized in Figure 7.19.

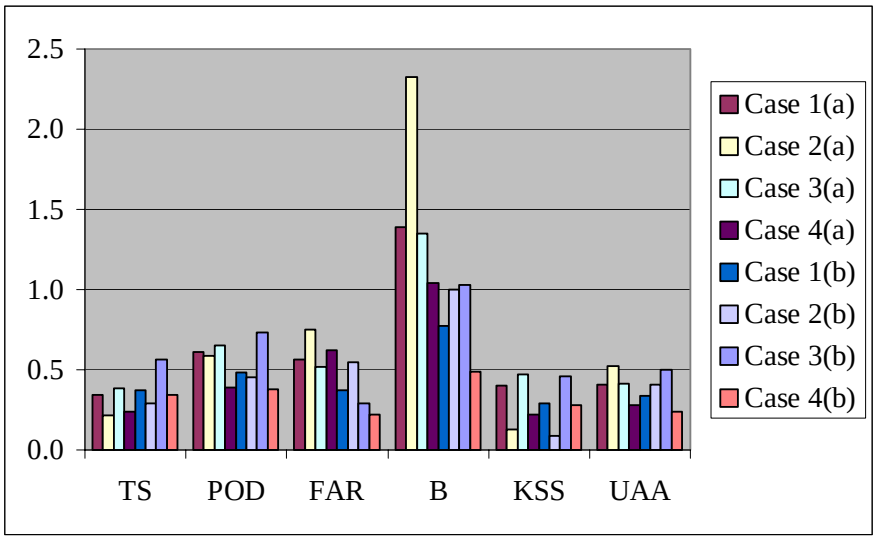

Figure 7.19: Forecast verification of classification tree analysis. Fidelity/Adamants.

The best threat score (TS) was achieved for the dataset combination of meteorological data from the Adamants and snowpack properties from Mt. Fidelity (Case 3) in both the learning and the test sample indicating that unstable days are best predicted in Case 3. However the scores with 0.38 and 0.56 in the learning and test sample respectively are not particularly high; a score of 0 indicates that none of the unstable days are correctly predicted. However this result is influenced by the unbalanced dataset (21% are DSI = 1) and the score on its own does not say much about the absolute performance, but allows a comparison between the cases. The higher score in the test sample is likely because the test sample is not as unbalanced (40% are DSI = 1). The TS is very low in Case 4, but even lower in Case 2 suggesting that unstable days are least accurately predicted when using the meteorological data from Mt. Fidelity, but none of the snowpack properties. In Case 1 the TS in the learning sample is with 0.34 similar to the score in Case 3 (0.38), but in the test sample with 0.38 compared to 0.56 less accurate.

The highest scores for the probability of detection (POD) were calculated for Case 3 in the learning and test sample indicating that false stable predictions were lowest in this dataset. As with TS, the POD for Case 1 in the learning sample is similar to the one in Case 3, but again the test sample in Case 1 is not as good as in Case 3. This suggests that the last two winters were not as well predicted, as is expected since the model was fitted to the learning sample.

In Case 2 and Case 4 the false alarm with 0.75 and 0.63 respectively were high in the learning sample, but better in the test sample (0.55 and 0.22 respectively). Again the unbalanced dataset might have influenced the results significantly. However in Case 3 false unstable days are less often predicted in both the test and learning sample compared to the other cases. The score is not high in the learning sample (0.52), but considering that the false unstable predictions have lower consequences than false stable prediction the result is reasonable, especially because Case 3 yielded the highest POD.

The bias in all of the learning samples indicates that unstable days are more often predicted than they actually occur. Case 2(a) has the highest overforecast with a B value of 2.33. In the test sample Case 1 and 4 underforecast and Case 2 and Case 3 predict approximately as many unstable days as often as they occurred. However comparing the scores with the results in Table 7.11 shows that in Case 2(b) half of the unstable days were misclassified and therefore the good bias is not as meaningful for determining whether unstable days were correctly classified.

The results for Case 2 indicate a random forecast in both the learning and the test sample, which is reflected in the Kuipers Skill Scores (KSS) close to zero (0.13 and 0.09 respectively). Again the best score was computed for Cases 3a and 3b (0.47 and 0.46). The KSS for Case 1(a) was reasonable (0.4) but not equally good as in Case 1(b).

Case 3, the meteorological data from the Adamants in combination with the snowpack properties from Mt. Fidelity, offered the best prediction of DSI considering the accuracy measures and the skill score and the results from the global cross-validation. However the forecast using only Mt. Fidelity data was reasonable showing that the meteorological data from an area farther away, but an elevation commonly skied has its

predictive value. All in all the Sk38, Load and H improved the forecasts significantly using classification tree analysis.

7.4.3 Nearest Neighbour analysis

Cornice was configured for each of the four cases used in the classification tree analysis. Table 7.12 summarizes the performance of the learning sample (a), the test sample (b) and the combined samples (c). The learning sample was used to configure Cornice and in (b) and (c) the same scales and weights were used in the forecast.

Table 7.12: Performance of Nearest Neighbour model. Fidelity/Adamants.

		Days forecast	**Unstable days (DSI = 1)**			**Stable days (DSI = 0)**			**Fitness of batch test (%)**
			Correct correct	**Wrong**	**%**	**Correct correct**	**Wrong**	**%**	
Case 1	(a)	207	26	8	77	139	34	80	78
	(b)	93	28	8	78	36	21	63	71
	(c)	300	56	14	80	174	56	76	78
Case 2	(a)	207	17	17	50	154	19	89	70
	(b)	93	18	18	50	43	14	75	63
	(c)	300	39	31	56	178	52	77	67
Case 3	(a)	207	27	7	79	151	22	87	83
	(b)	93	21	15	58	49	8	86	72
	(c)	300	45	25	64	197	33	86	75
Case 4	(a)	207	24	10	71	133	40	77	74
	(b)	93	20	16	56	32	25	56	56
	(c)	300	45	25	64	148	82	64	64

Similar to the observations in the classification tree analysis the incorporation of Sk38, Load and H enabled a better forecast of skier-triggered dry slab avalanches on persistent weak layers in the Columbia Mountains (Table 7.11). In Case 1 compared to Case 2 the forecast improved by 8%, 8% and 11% in the learning sample (a), the test sample (b) and the combined samples (c) respectively and in Case 3 compared to Case 4 by 9%, 16% and 11%. However the amount of correctly forecast unstable days improved in Case 1 by 27%, 28% and 24% and in Case 3 by 8%, 2% and 0%, implying that the three

snowpack properties greatly improved the forecast for unstable days when using the Mt. Fidelity data and in the Adamants dataset the stable days also were improved.

The best forecast for unstable days was made in Case 3 with 79% correct forecasts, though the forecast in the test sample at 58% was not as good. The test sample was best predicted in Case 1. Interesting to notice is that Case 1(c) predicted unstable days with 80% accuracy, though the learning and the test sample was only predicted with 77% and 78% respectively. This suggests that by the inclusion of the test sample more nearest neighbours similar to the once in the learning sample were added and consequently a better forecast in the learning sample was possible. The same is observed for Case 2 but not Case 3 and Case 4. The difference must be because of the meteorological data recorded in the Adamants. The reason can either be that the last two years were different from the other years or the measurements were different, for instance, by a change in the study plot location.

These results are different from the classification tree results in Section 7.4.2 and are reflected in the accuracy scores and skill scores (Figure 7.20).

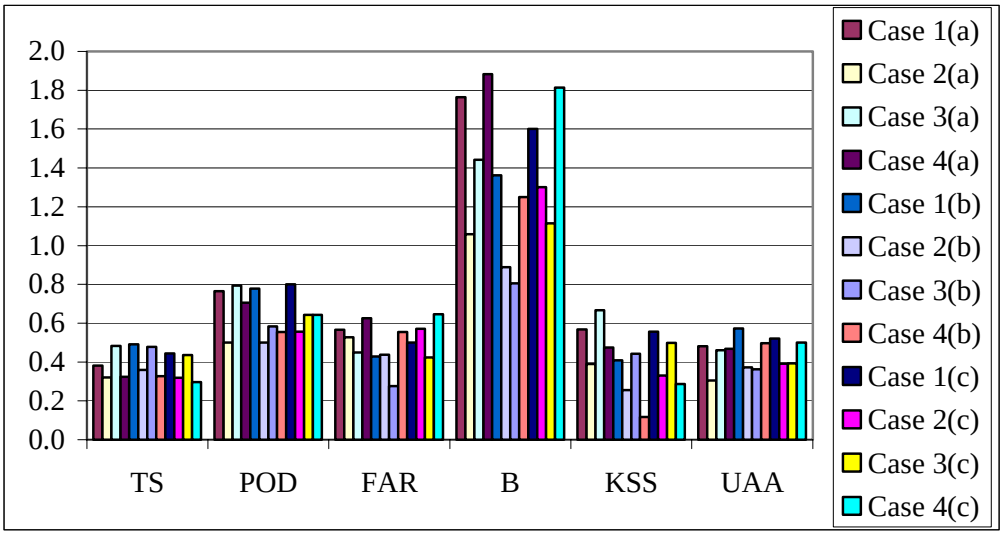

Figure 7.20: Forecast verification of nearest neighbour analysis. Fidelity/Adamants.

In the forecast verification of the results from the classification tree, the scores indicate that Case 3 offered the best forecast results. The threat score (TS) calculated from

the results of the nearest neighbour analysis is still best in the learning sample in Case 3(a) (0.48) and therewith better than for Case 1(a) (0.38). The TS is similar in the test samples: 0.49 in Case 1(b) and 0.48 in Case 3(b). The absolute scores are still poor and indicate that less than 50% of the unstable days were correctly predicted, which is a result of the unbalanced dataset.

The false stable predictions were lowest in Case 3(a) (highest POD score), though in Case 3b these predictions were higher than in Case 1(b) (0.58 compared to 0.64) also indicating that the test sample is better predicted with the meteorological data from Mt. Fidelity.

The false alarm (FAR) was lowest in the learning sample (0.45) and the test sample (0.28) of Case 3; FAR was highest in Case 4(a) and (b). The false alarm was lower in the test sample of Case 1 compared to the learning sample of Case 1 with 0.43 and 0.57. The lowest false alarm rate was observed in Case 3(b) with 0.28, though the POD was not good and therefore often predicting false stable conditions.

The bias indicates an overforecast in all learning and combined samples. Case 2(a) had a calculated bias of 1.06, indicating an only slight overforecast and almost as many days were forecast as observed. However 50% of the unstable days were misclassified. Generally an overforecast is more desirable than an underforecast, because the chance of false stable predictions is lower.

The Kuipers skill scores (KSS) is low for the datasets excluding Sk38, Load and H and highest in Case 3(a) and Case 3(b), but in the combined dataset better in Case 1(c), again reflecting that the test sample added similar days to the dataset in Case 1.

The results from the classification tree and the nearest neighbour analysis differ from each other. In the classification tree Case 3 offered the best predictions in both the learning and the test samples. However the results of the nearest neighbour analysis showed a better classification of the learning sample in Case 3 though the test sample was better predicted in Case 1. All the calculated scores were better for the cases in which Sk38, Load and H were included in both the classification tree and the nearest neighbour analysis. Overall, inclusion of snowpack properties improved the forecast of skier-triggered avalanches. A clear advantage of using the snowpack properties from Mt. Fidelity to

predict the avalanche activity in the Adamants was shown, however there is no apparent advantage of using the meteorological data from Mt. Fidelity at 1905 m compared to the meteorological data from the Adamants at 1100 m even though the higher elevation is commonly more skied.

7. 5 Summary

The main hypothesis of this thesis is that the incorporation of stability indices and snowpack properties (Sk38, H and Load) improve the forecast of skier-triggered avalanches on persistent weak layers.

In this section a skier instability index DSI was introduced as response variable to allow the inclusion of days on which skier-triggered avalanches were likely but due to limited skiing or terrain choices no avalanches were skier-triggered. The index is based on actual avalanche activity and the end-of-the-day stability ratings.

In total 26 potential predictor variables were assessed in regard to their value to forecast DSI. The results of the Blue River study area showed that Sk38 and Load have high predictive potential whereas the correlation with *H* was not as significant in the Spearman rank correlation.

Classification tree analysis showed that Sk38 and Load improved the forecast whereas H contributed only marginally to the result.

Including all variables into a Nearest Neighbour forecasting model yielded the most accurate forecast for skier-triggered avalanches on persistent weak layers in the area around Blue River with a forecast accuracy of 72%-78% compared to 64%-66% when excluding Sk38, Load and *H*. However it was also shown that the misclassification of avalanche days is a result of missing information about the snowpack, for instance a thick crust above a weak layer may prevent avalanche releases even though the shear strength of the weak layer is low.

Cornice, the Neareast Neighbour Software used in this study, weights the predictor variables automatically to optimize the models output. However it is thought that the weights set by a forecaster might improve the forecast due to the local knowledge. Even though the forecast for the last two years improved (82%) the prior year were only forecast

with 63%. The results show that the weights set by the forecasters have their value, but the optimization in Cornice is more advanced and finds the best statistical fit.

The inclusion of results from rutschblock tests that were done throughout the study area into a regional forecasting model did not improve the forecast and raises the question whether point measurements are valuable in a forecasting model on a regional scale.

In the second study DSI was based on avalanche activity data and stability evaluations in the Adamants. The meteorological predictor variables were available from a study plot at 1100 m in the Adamants close to the ski runs but also from a weather station on Mt. Fidelity about 50 km apart from the ski runs but at an elevation commonly skied (1905 m). In addition snowpack properties were measured at Mt. Fidelity. It has been shown that this Sk38, Load and *H* improved the forecast when using the Adamants weather data and when using the Mt. Fidelity weather data. However is has also been shown that there is no obvious advantage of using the weather data from an elevation commonly skied instead of using the data from a study plot near by the ski runs. The forecast accuracy with Sk38, Load and *H* yielded fits of 71%-83% and 56%-70% without these snowpack properties.

In this chapter it has been proven that a skier stability index and snowpack properties improve the regional forecast for skier-triggered dry slab avalanches on persistent weak layers.

8. Forecasting skier-triggered avalanches on non-persistent weak layers

8.1 Introduction

As previously mentioned, the threat to skiers from avalanches triggered on non-persistent weak layers, including storm snow instabilities, is lower than from avalanches on persistent weak layers. In addition, the poor stability during storms is often more obvious and therefore easier to forecast. However, a daily Avalanche Forecasting Model is not complete without the prediction of storm snow avalanches. As pointed out in Section 2.2.1, storm snow avalanches may be considered as releases on a weak layer of persistent or non-persistent crystal forms during a storm, or only as releases on non-persistent weak layers. In the second case, avalanches on surface hoar or facets, for example, would not be considered as a storm snow avalanche during the first storm that buries the persistent weak layer. In this chapter the skier stability index Sk38 is analyzed to assess the predictive potential considering storm instabilities involving both persistent and non-persistent weak layers.

8.2 Datasets

The first dataset is a combination of data from Glacier National Park and the Adamants (Figure 3.1, Chapter 3) and is used to assess the predictive value of shear frame data to forecast skier-triggered dry slab avalanches in storm snow. The second dataset is from the area around Blue River and is used to assess the forecasting model without using shear frame data. Both study areas were already used in Chapter 7 for persistent weak layers.

Meteorological and shear frame data from Mt. Fidelity were used to predict the avalanche activity in the Adamants, an area about 50 km farther north. The Mt. Fidelity weather and study plot is located at 1905 m and therefore is at an elevation commonly skied. A study plot in the Adamants is located at 1100 m, an elevation less relevant to skier triggering, but geographically closer to the ski runs. This approach presumes that the snowpack and weather conditions are similar over this region of the Columbia Mountains so that the Mt. Fidelity data is relevant to the avalanche conditions in the Adamants. This

has been shown for persistent weak layers in Chapter 7, where the performance of a nearest neighbour model for both study plots, using Mt. Fidelity weather and snowpack data instead of the weather data available in the Adamants was compared. As mentioned above storm snow avalanches may release on persistent or non-persistent crystal form and the definition of avalanches on storm snow instabilities can involve persistent weak layers. To assess whether a model for storm snow instabilities or persistent weak layers better predicts avalanches which release during the first storm that buries a weak layer the steps of analysis involved in this thesis are:

1) configuration of a forecasting model on persistent weak layers including avalanches occurring during the first storm that buries them (Chapter 7),
2) configuration of a forecasting model including all avalanches which release during or soon after a storm (this chapter),
3) configuration of a forecasting model including all avalanches which occurred during or soon after a storm and any avalanche which released on a non-persistent layer between storms (this chapter), and

8.3 Methods

As described in Chapter 7, incorporating a stability index and snowpack properties improved the forecast of skier-triggered avalanches on persistent weak layers. Based on this idea, a similar approach to forecast skier-triggered avalanches on storm snow instabilities seemed to be worthwhile.

The avalanche control section in Glacier National Park uses a stability index as the basis to determine the avalanche hazard of storm snow layers for natural avalanches that might threaten the highway (Schleiss and Schleiss, 1970). For this purpose they measure the shear strength of weak layers with a 0.01 m^2 shear frame and calculate a simple stability ratio, SF as described in Section 2.4.3. Schleiss and Schleiss (1970) suggest that a calculated SF of 1.5 or less indicates critical stability for natural avalanches. The question in this chapter was whether the same shear frame measurements could be used to forecast skier-triggered avalanches rather than natural avalanches as will be discussed in Chapter 9.

For this purpose the measured shear strength value was adjusted for the shear frame size and the normal load (Jamieson and Johnston, 1998). Further, instead of calculating SF, a skier stability index called Sk38 was calculated, which accounts for the ski penetration and slope angle (38°) typical of avalanche starting zones (Section 2.4.3).

During the attempts to predict the shear strength changes of storm snow instabilities over time without pulling shear frames a few problems were encountered. These included:

1. multiple weak layers in the storm snow, which can be deposited within hours
2. shear frames were generally pulled once a day whereas stability changes of some weak layers occur within hours
3. difficulty tracking specific layers for two or more days based only on their depth adjusted for snowfall and settlement.

As a consequence of these problems in the dataset, an empirical model to calculate shear strength changes over time could not be found. Instead, *measured* shear strength values were used to assess the predictive value of stability indices for skier-triggered avalanche activity on storm snow instabilities. This is in contrast to the analysis of persistent weak layers, where models were found to calculate the shear strength of a weak layer over time. However, for storm snow instabilities it was possible to use the measured shear strength values, because these were measured almost daily in the Mt. Fidelity study plot during storms and shortly after. In addition, a simple approach will be used, which is described later, to assign values to Sk38 in order to be able to run a daily forecasting model.

The complete dataset included nine winters (1995/1996 – 2003/2004), each starting on December 1 and ending March 31, totaling 1092 forecasting days. Shear frame measurement were performed at Mt. Fidelity on 530 of these days. CMH Adamants provided the avalanche occurrence data, but because their season does not usually start up before the middle or end of December, the dataset is reduced accordingly.

8.4 Response variable

Similar to the analysis on persistent weak layers in Chapter 7, a skier instability index (DSI) was defined as a response variable. DSI was assigned a value of one for each day in which one or more skier-triggered dry slab avalanches on storm instabilities were reported, or the ski guides operating in the area rated the stability at the end of the day as poor or very poor (Canadian Avalanche Association, 2002). Fair-poor ratings were not reported in this study area on non-persistent weak layers. In this analysis avalanches with a minimum size of one were considered (Canadian Avalanche Association, 2002). All other days were DSI = 0. This index allows the inclusion of days on which skier-triggered avalanches were likely, but due to limited conditions for flying (hence limited helicopter skiing) or the terrain selection of the guides, no dry slab avalanches on storm snow instabilities were skier-triggered.

A daily skier instability index (DSI) was defined for 889 out of 1092 days with 624 stable days (DSI = 0) and 265 unstable days (DSI = 1); actual avalanches were reported on only 157 days. Of these, most avalanches released on surface hoar layers (61) and on layers of faceted crystals (6), while the remainder were on unidentified weak layers within the storm snow. In the analysis for forecasting skier-triggered avalanches on non-persistent weak layers seven more days were defined as DSI = 1 where an avalanche occurred on a non-persistent weak layer between storms.

In addition to forecasting DSI, an attempt was made to forecast the maximum size of an avalanche occurrence. The statistical analysis showed no significant correlations, possibly because avalanches only occurred on 157 days.

8.5 Predictor variables

Table 8.1 summarizes the daily predictor variables used in the statistical analysis for both study areas.

Table 8.1: Daily predictor variables

Abb.	Definition	Blue River	Rogers Pass
T	Air temperature at a.m. (°C)	✓	✓
WD	Wind direction at a.m.; (minus 90° modulo 360) east as base azimuth (°)	✓	✓
RH	Relative humidity at a.m. (%)	✓	✓
WS	Wind speed at a.m. (km h^{-1})	✓	✓
WrunY	24 h wind run for previous day (km)	✓	✓
HNY	Height of new snowfall for previous 24 h (m)	✓	✓
HS	Height of snowpack at a.m. (m)	✓	✓
PcpY	Water equivalent of precipitation on previous day (mm)	✓	✓
Strm	Cumulative new snowfall (storm) since last day with less than 0.3 mm of precipitation (m)	✓	✓
Strm3day	Cumulative new snowfall (storm) over the last three days (m)	✓	✓
TminY	Minimum temperature for previous day (°C)	✓	✓
TmaxY	Maximum temperature for previous day (°C)	✓	✓
RHmnY	Minimum relative humidity for previous day (%)	✓	✓
RHmxY	Maximum relative humidity for previous day (%)	✓	✓
WV	Wind vector, magnitude of the direction and intensity of the wind at a.m.	~	✓
WSavg	Average wind speed over 24 hours	~	✓
WDavg	Average wind direction over 24 hours; (minus 90° modulo 360) east as base azimuth (°)	~	✓
WVavg	Average wind vector over 24 hours	~	✓
TriPrev	Number of skier-triggered dry slab avalanches on previous day	~	✓
TriPrevBR	Skier-triggered avalanches in storm snow on previous day (yes/no)	✓	~
Sk38	Skier stability index	~	✓

The variables were chosen based on their availability and consistency.

In this analysis, the skier stability index Sk38 was calculated, using measured shear strength values with a shear frame. This index includes the calculated ski penetration, slope angle and load of a skier. The index cannot be calculated when the ski penetration is deeper than the weak layer (Section 2.4.4). However, Sk38 was assigned a value of 10 when the calculated ski penetration exceeded the weak layer depth because the storm snow was most likely not hazardous to skiers and small avalanches are not likely reported consistently. Even though avalanches might release due to the up and down motion of skiing if the ski penetration is deeper than the weak layer this fact is disregarded the resulting avalanches

would involve a relatively thin slab of soft snow that is rarely a threat to a skier. Altogether shear frames were pulled on 530 days. Of these, ski penetration did not exceed the depth of the weak layers on 184 days and consequently Sk38 was assigned a value of 10 for the remaining 346 days. These 346 also include days on which testing was attempted but no shear fractures occurred. In a second approach, all days on which a DSI could be defined were considered to evaluate the predictive potential of Sk38 in a daily forecasting model to forecast avalanches on non-persistent weak layers. Shear strength measurements are not available daily and therefore a simple approach was chosen to indicate stable conditions by assigning a value of 10 to days on which shears in storm snow were unlikely and therefore not tested, as already done for days on which the ski penetration was deeper than the slab depth. In the remainder of this chapter Sk38+ will indicate the stability index in the second approach. This led to the consideration of two main cases:

Case 1:	Only days with shear frame measurements during and soon after a storm including Sk38
Case 2:	Only days with shear frame measurements during and soon after a storm excluding Sk38
Case 3:	All days on which a DSI could be defined including Sk38+
Case 4:	All days on which a DSI could be defined excluding Sk38+

8.6 Statistical methods

The statistical methods are similar to those used in Chapter 7. Again Spearman rank correlations are used to assess the importance of each predictor variable on the response variable DSI, then classification tree analysis is applied to determine the importance of the combined relations of the predictor variables with DSI and last a Nearest Neighbour model is configured to assess the performance of a model which is used in avalanche forecasting operations. Because of the limited availability of the response variable DSI and the predictor variable Sk38 in the classification tree analysis only 400 days were considered in Case 1 and 2, and 823 days for Case 3 and 4. Table 8.2 summarizes the samples of both cases.

Table 8.2: Summary of dataset Rogers Pass/Adamants.						
	Case 1 + 2			Case 3 + 4		
	N	DSI (days) 1	0	N	DSI (days) 1	0
Learning sample	330	143	187	651	193	458
Test sample	70	33	37	172	55	117
Total	400	176	224	823	248	575

The test sample consists of data observed during the past two winters.

8.7 Analysis

8.7.1 Spearman rank correlations with Daily Instability Index (DSI)

The results from the Spearman rank correlations (Case 3) are summarized in Table 8.3.

Table 8.3: Spearman rank correlations with DSI. Insignificant correlations are italics (p > 0.05).

Variable	Valid N	Spearman R	p-level
PcpY	852	0.366	<1E-17
Strm3day	852	0.357	<1E-17
HNY	881	0.300	<1E-17
Strm	848	0.295	<1E-17
RHminY	881	0.263	2.1E-15
Sk38+	889	-0.260	3.3E-15
RH a.m.	880	0.222	2.5E-11
RHmaxY	881	0.206	6.7E-10
WVavg	889	0.176	1.2E-7
WSavg	889	0.172	2.5E-7
WrunY	889	0.172	2.5E-7
T a.m.	879	0.168	5.2E-7
TriPrev	871	0.165	9.5E-7
TminY	881	0.139	3.4E-5
WV	883	0.125	2.1E-4
WS a.m.	889	0.124	2.2E-4
SK38	430	-0.166	5.3E-4
WDavg	889	-0.101	0.003
WD a.m.	883	-0.087	0.009
TmaxY	881	0.082	0.015
HS a.m.	*889*	*0.016*	*0.643*

Most of the predictor variables correlated with DSI, except the height of the snowpack and the wind direction measured in the morning. The explanations for the significant predictor variables are as follows:

Temperature: The minimum, maximum and the temperature in the morning showed positive correlations, indicating that warmer temperatures are associated with unstable days. This is likely because storms are generally associated with warmer temperatures.

Relative humidity: The minimum, maximum and the humidity in the morning are correlated positively with DSI showing that snow storm avalanches are associated with high humidity, which is not surprising because high humidity is generally observed during storms when warm, moist air is present (McClung and Schaerer, 1993: 161).

Precipitation, new snow (PcpY, HNY, strm, strm3day): The positive correlations indicate that the stress applied to the snowpack due to loading contributes to avalanches and the more rapidly load is applied the more the shear strength of weak layers will lag behind the stress (Schweizer et al., 2003). PcpY and strm3day are stronger correlated than HNY probably because they are more directly related to the load because they are measures of the water equivalent over the past 24 hours and three days respectively.

Wind: The positive correlations between avalanche days and wind speed variables are expected because at higher wind speed there is often greater loading of start zones.

Skier-triggered avalanche activity from previous day (TriPrev): The positive correlations indicate that when more avalanches were triggered the day before, the more likely the forecast day is an avalanche day. This is expected because storms often last longer than one day.

Sk38: The negative correlation suggests that lower values of Sk38 are associated with avalanche days, as expected.

The Spearman rank correlations for Case 1 showed similar correlations except that the temperature was correlated more strongly, likely because during and soon after a storm an increase in temperature contributes to instability (Schweizer et al., 2003).

Compared to the Spearman rank correlations for skier-triggered avalanches on persistent weak layers (Chapter 7) in the Blue River dataset T, TminY and TmaxY are significantly and positively correlated with DSI. This is likely because a during storm the temperature becomes more relevant and changes in temperature affect surface layers more than deeper layers (Schweizer et al. 2003).

8.7.2 Tree analysis to predict Daily Instability Index (DSI)

The first classification tree was constructed for Case 1 including Sk38 as a predictor variable. The best classification for avalanche days was obtained when setting the prior probabilities to 0.55 for unstable days and 0.45 for stable days. In the classification tree, the algorithm only selected precipitation as a splitting variable, as shown in Figure 8.1.

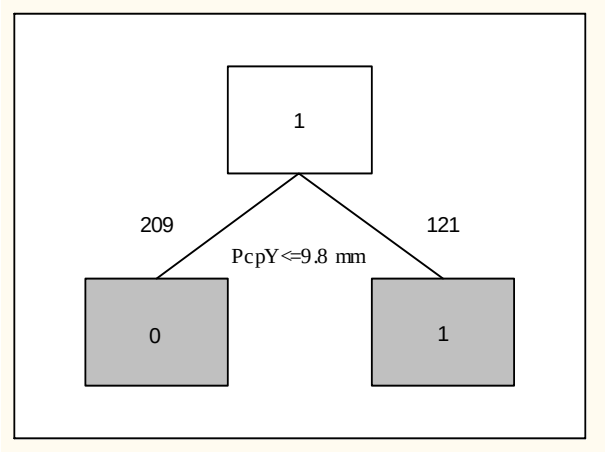

Figure 8.1: Classification tree. Rogers Pass study area (N=330). Case 1(a).

The critical value for the split is 9.84 mm implying that with lower precipitation over 24 hrs, avalanches are less likely (left branch). The global cross-validation results are shown in Table 8.4.

Table 8.4. Global cross-validation results from classification tree. The learning sample (L) consists of 143 unstable days and 187 stable days. The test sample (T) consists of 33 unstable days and 37 stable days.

		Misclassification						
Variables	**Sample**	**Unstable days (DSI = 1)**		**Stable days (DSI = 0)**				**Priors DSI=0/ DSI=1**
		No	**(%)**	**No**	**(%)**	**Cost**	**Error**	
Case 1 and 2	L	45	(32%)	92	(49%)	0.395	0.027	0.45/0.55
	T	16	(49%)	12	(32%)	0.413	0.059	

In the learning sample, the unstable days were forecast correctly in 68% of the cases, and in only 51% of the cases for stable days. The fit of the test sample is exactly the opposite with 68% correctly classified stable days and 51% correctly classified unstable days. The calculated misclassification cost is with 0.395 lower in the learning than in the test sample with 0.413, which is a result of the set prior probabilities. Sk38 did not play a role in the development of the classification tree construction and excluding Sk38 from the set of predictor variables did not influence the results. Consequently Case 1 and 2 are further discussed together.

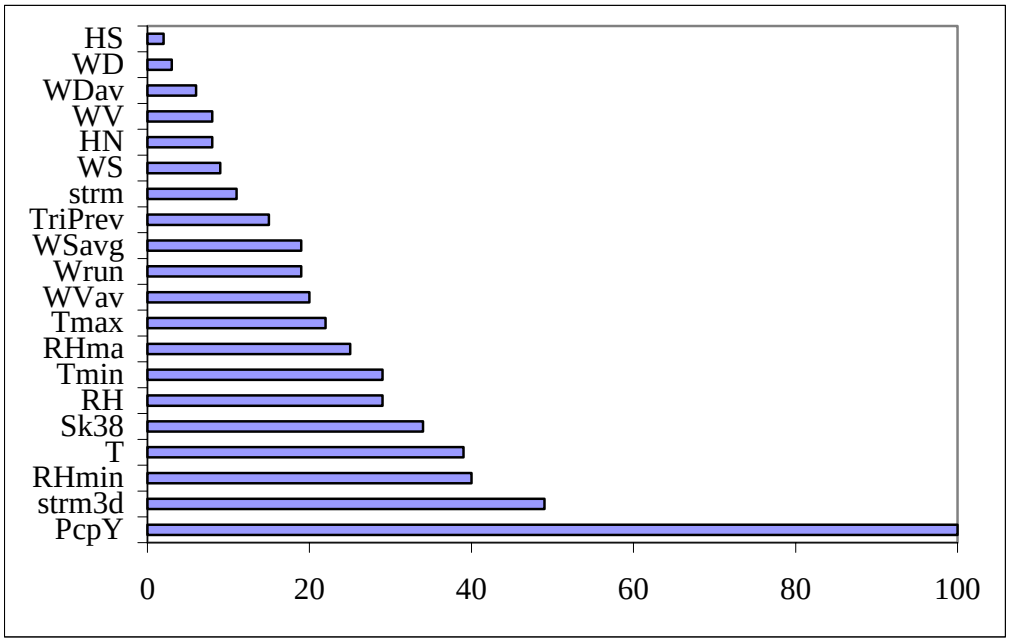

Figure 8.2: Importance ranking. Case 1

PcpY dominates the importance ranking with a value of 100 on a scale from 0 to 100 followed by strm3day, RHmin, T and Sk38 out of the 15 predictors (Figure 8.2). However these variables do not play a role in the classification tree development with the chosen stopping conditions. Excluding Sk38 did not affect the results.

Because the classification tree algorithm did not select Sk38 and the results are dominated by the precipitation over the past 24 hours the skier stability index did not seem to add value to forecast storm snow instabilities which might be because the stability of storm snow instabilities often change within hours but shear frame measurements are usually done once a day. However, in a second approach (Case 3 and 4) a bigger dataset will be analyzed to assess the value of Sk38+ on a daily basis to predict storm snow instabilities and to provide a forecasting model, which can then be used in combination with the Forecasting Model for skier-triggered avalanches on persistent weak layers. In total 823 days were considered. The best priors to maximize the correctly forecast avalanche days were 0.4 for DSI = 0 and 0.6 for DSI = 1, which accounts for the unbalanced dataset and the higher consequences of incorrectly classifying an avalanche day as a non-avalanche day. Figure 8.3 shows the resulting classification tree.

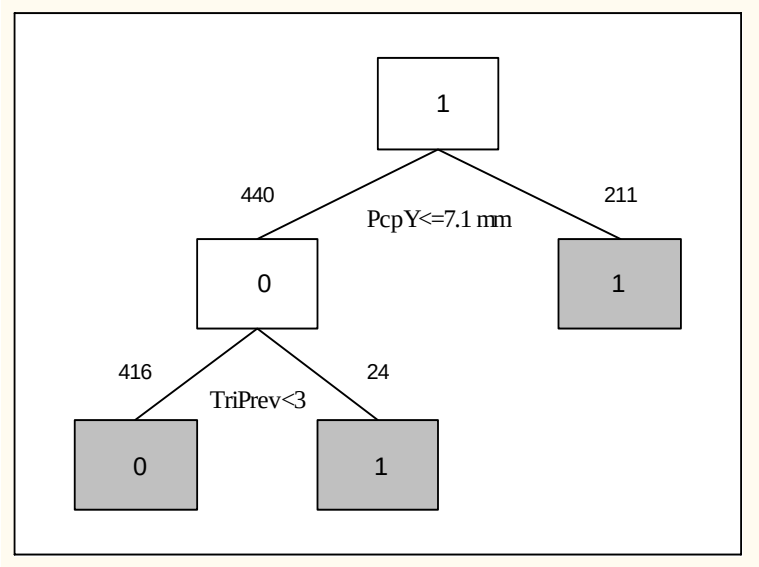

Figure 8.3: Classification tree. Rogers Pass study area, with SK38+ (N=651)

The classification tree consists of two splits with two splitting variables. Again the precipitation over the past 24 hours was selected as the first splitting variable and thereby contributed most to the fit of the tree. The right branch of the tree predicts avalanches when the PcpY was greater than 7.1 mm over the past 24 hours. In the left branch, the next splitting variable was the number of triggered avalanches on storm snow on the previous day. In the case that four or more avalanches were skier-triggered on storm snow on the previous day, an avalanche is more likely on the forecast day.

The misclassification in the model building was 67 of 193 (35%) for the unstable days and 109 of 458 (24%) for the stable days. This means that 65% of the unstable days and 76% of the stable days were classified correctly. The global cross-validation results in Table 8.5 show a similar result in the learning sample with 67% correctly forecast unstable days and 71% correctly forecast stable days. The misclassification cost equals 0.316. However, the performance of the test sample was less accurate with only 53% correctly forecast unstable days and 81% correct stable days. As a consequence of the set priors, the misclassification cost was higher in the test sample than in the learning sample (0.359 compared to 0.316).

In the importance ranking (Figure 8.4) PcpY was ranked highest, as observed in Case 1, underlining the importance of this predictor variable. However four variables gained substantially in importance: strm3day, TriPrev, Sk38+ and Rhmin, although only PcpY and TriPrev were used in the tree development with the selected splitting conditions.

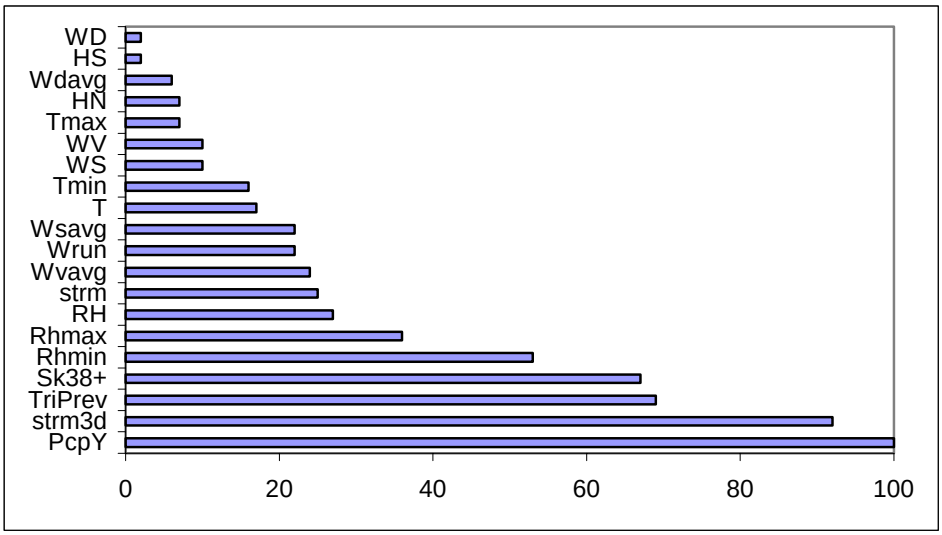

Figure 8.4: Importance ranking (with Sk38+)

The tree in Figure 8.5 was developed when Sk38+ was excluded from the analysis (Case 4). The first two splits were identical with the splits in Case 3, but in a third split avalanches were determined when strm3day exceeded 23.7 mm.

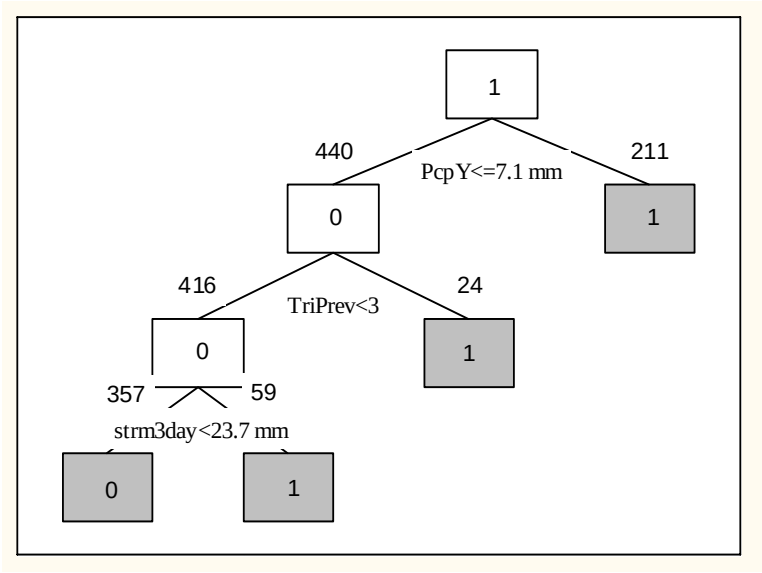

Figure 8.5: Classification tree. Rogers Pass study area, without Sk38

The misclassification was 48 of 193 (25%) for unstable days and 149 of 458 (33%) for non-avalanche days. This means that 75% of the unstable days were correctly forecast, which is a better result compared to Case 3.

Table 8.5 summarizes the global cross-validation results of the tree analysis for Cases 3 and 4.

Table 8.5. Global cross-validation results from classification tree. The learning sample consists of 193 unstable days and 458 stable days. The test sample consists of 55 unstable days and 117 stable days.

		Misclassification						
Variables	**Sample**	**Unstable days (DSI = 1)**		**Stable days (DSI = 0)**				**Priors DSI=0/ DSI=1**
		No	**(%)**	**No**	**(%)**	**Cost**	**Error**	
Case 3	L	64	(33%)	134	(29%)	0.316	0.022	0.4/0.6
	T	26	(47%)	22	(19%)	0.359	0.043	
Case 4	L	63	(33%)	127	(28%)	0.307	0.022	0.4/0.6
	T	24	(44%)	35	(30%)	0.382	0.044	

The misclassification cost was reduced in Case 3 compared to Case 4 from 0.316 to 0.307 in the learning sample, but increased in the test sample from 0.359 to 0.382. However, comparing the performance of the unstable days in the learning sample Case 4 better predicted avalanches by one case and in the test sample by two cases. The higher cost in the test sample in Case 4 compared to Case 3 is a result of the higher misclassification of stable days.

The importance ranking (Figure 8.6) for Case 4 shows that the cumulative new snowfall (storm) over the past three days and the precipitation were the most important factors for forecasting skier instability (DSI) in the Rogers Pass dataset, followed by the number of skier-triggered avalanches on the previous day and RHmin. The only important difference to the ranking in Case 3 is that the strm3day is ranked first place and precipitation in second.

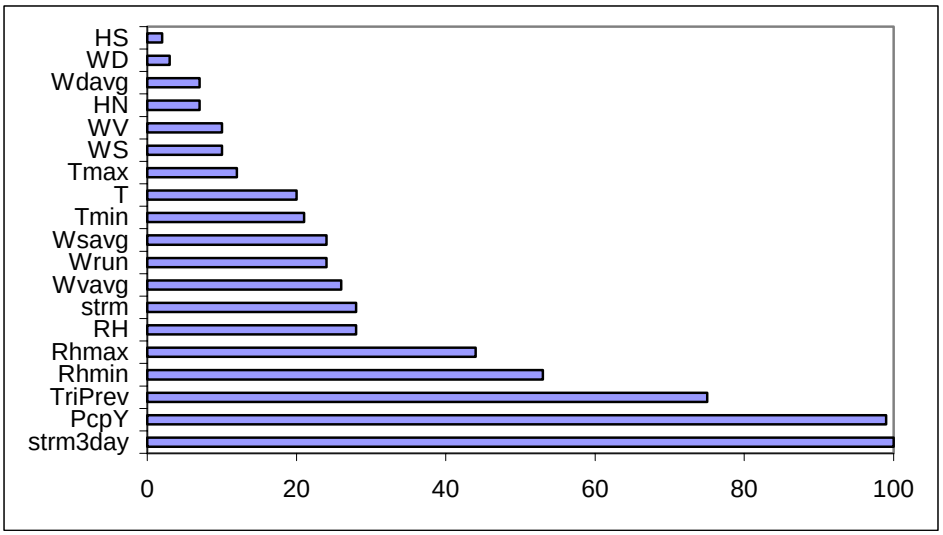

Figure 8.6: Importance ranking of predictors (without Sk38+)

The importance of cumulative storm snow indices was recognized in earlier studies (Perla 1970).

In conclusion, Sk38 did not improve the forecast in an obvious way, using the simple approach in this study for defining stable days. A shear strength model for storm snow instabilities could not be defined and consequently the use of Sk38 requires manual shear strength measurements, which are time consuming and usually not performed in forecasting operations. However, the results are promising and likely to improve when a more refined shear strength definition becomes available based on more consistent shear strength measurements during and in between storms.

The forecast verification in Figure 8.7 confirms the interpretation that Sk38 did not improve the forecast in an obvious way.

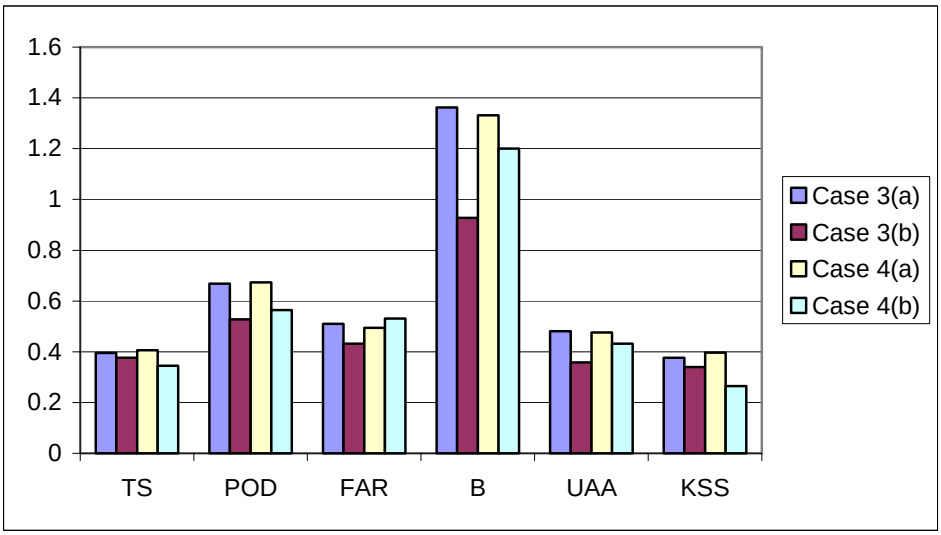

Figure 8.7: Forecast verification. Classification tree analysis. Rogers Pass/Adamants.

In all the accuracy measures and skill scores the results in the learning samples of Case 3 and 4 are similar where the results for Case 4 indicate a marginally better forecast. For instance the threat score (TS) is calculated with 0.39 in Case 3 and 0.41 in Case 4 indicating that unstable days are a little better forecast without using Sk38 as predictor. However the unweighted average accuracy (UAA) in Case 4(b) is higher than in Case 3(b) indicating that the unweighted mean of the model accuracy at forecasting unstable and stable days is higher, which accounts better for rare events. False stable predictions were forecast less often in Case 4(a) than in Case 3(a) with a probability of detection (POD) of 0.68 compared to 0.67, again an only marginally better score. A similar result can be observed for the false alarm (FAR) with a lower and therewith better score for Case 4(a). The bias (B) indicates an overforecast in both learning samples. The Kuipers skill score (KSS) is also quite similar in Case 3(a) and 4(a) with 0.38 compared to 0.4. All the scores are low as a result of the unbalanced dataset, which is also reflected in the test samples. However the test sample in Case 4 performs less accurately than the Case 3 indicating that the last two years were better forecast using Sk38 as predictor variable based on higher scores for TS, POD, UAA and KSS and a lower score for FAR. However B calculated an underforecast for Case 3(b), which might be a concern in avalanche forecasting because

less unstable days were predicted than occurred and therefore the probability of false stable forecasts is higher than in a bias that indicates an overforecast of unstable days.

All in all the results showed that Sk38 in its current format added little or no predictive value to the forecast, though the last two years seemed to be better forecast in the classification tree analysis. However Sk38 might be a useful predictor if shear strength measurements for storm snow instabilities become more available. Anyway, both Cases will be configured in Cornice to assess the performance in an avalanche forecasting model.

8.7.3 Nearest neighbour analysis

Cornice was configured for each of the two cases used in the classification tree analysis. Table 8.6 summarizes the performance of the learning sample (a), the test sample (b) and the combined samples (c). The learning sample was used to configure Cornice and in (b) and (c), the same scales and weights were used in the forecast.

Table 8.6. Performance of Nearest Neighbour model. Rogers Pass/Adamants.

		Days forecast	*Unstable days (DSI = 1)*			*Stable days (DSI = 0)*			*Fitness of batch test (%)*
			Correct	*Wrong*	*% correct*	*Correct*	*Wrong*	*% correct*	
Case 3	(a)	695	141	62	70	333	159	68	69
	(b)	194	37	25	60	91	41	69	65
	(c)	889	178	87	67	416	208	67	67
Case 4	(a)	695	142	61	70	309	183	63	66
	(b)	194	46	16	74	80	52	61	68
	(c)	889	185	80	70	383	241	61	66

The results from the nearest neighbour analysis are in some aspects different from the results in the classification tree analysis. The forecast of Cornice predicted unstable days more accurately than the classification tree in the learning and the test sample for both cases under consideration. The improvement for Case 3(a) was 3%, Case 3(b) 7%, Case 4(a) 3% and Case 4(b) 18%. However the classification in the tree analysis was better for stable days with 3% fewer misclassifications in Case 3(a), 12% in Case 3(b), 9% in Case 4(a) and 9% in Case 4(b). For avalanche forecasting purposes the results from the nearest neighbour analysis are preferable.

Comparing the performance of Cornice including and excluding Sk38 from the analysis shows that the overall performance in the learning sample is better in Case 3(a) with 69% compared to 66% in Case 4(a) though the forecast for unstable days equals 70% in both cases. However the forecast in the test sample is better without Sk38 with 68% compared to 65% for the batch forecast and 74% compared to 60% for the unstable days. The forecast for stable days in Case 3(b) it more accurate (69%) than in Case 4(b) (61%).

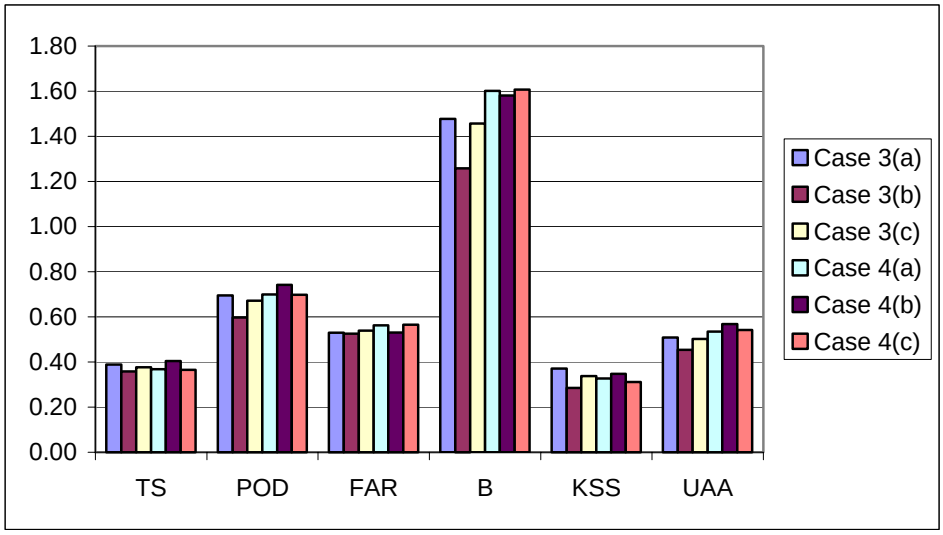

Figure 8.8: Forecast verification. Nearest neighbour analysis. Rogers Pass/Adamants.

The better forecast of Case 3(a) for unstable days is reflected in the accuracy and skill scores with higher TS, POD and KSS scores (Figure 8.8) and the lower FAR. In addition, the overforecast in B is not as accentuated as in Case 4(a). However the UAA is slightly lower. The performance of the test sample in Case 4(b) is significantly better than in Case 3(b) with higher scores for TS, POD, KSS and UAA, though the bias indicates a higher overforecast.

This analysis does not show that Sk38 (based on shear frame measurements) substantially improves the forecast for storm snow avalanches.

All in all it seems that the forecasting of skier-triggered avalanches on storm snow instabilities is quite difficult with the methods in this chapter and available data. As in

previous studies of natural avalanches, precipitation is the most important predictor. One possible reason might be that avalanches on storm snow instabilities do not occur very often due to limited skiing or they are often not large enough to be consistently recorded, which would imply that the defined skier instability index does not represent the instability but an index of hazard to a skier.

8.8 Results in Blue River

The dataset from Blue River consists of data from the winters of 1991/1992 to 2002/2003. The aim is to forecast dry slab avalanches on non-persistent weak layers and therefore time periods from the first of December to the 31 of March were considered. However in some years the data were not consistently available reducing the dataset accordingly. In total 1381 days were considered of which 302 (22%) were classified as unstable days (DSI = 1) and the remaining 1079 days as stable (DSI = 0); again an unbalanced dataset.

In the study area around Blue River the testing of *non-persistent* weak layers with a shear frame was only done on an irregular basis and therefore this section does not assess the predictive value of snowpack properties. It however evaluates the performance of a nearest neighbour model configured with meteorological data and the triggered avalanches on the previous days, to be able to run the model on a daily basis.

Similar to the analysis in previous chapters, the last two years were excluded from the entire dataset as a test sample. However in this analysis the last two years in the dataset are 2001/2002 and 2002/2003; the past winter could not be considered because of inconsistency in the recordings. Table 8.7 summarizes the performance of the learning sample (a), the test sample (b) and the combined samples (c).

Table 8.7. Performance of Nearest Neighbour model. Blue River.

		Days forecast	*Unstable days (DSI = 1)*			*Stable days (DSI = 0)*			*Fitness of batch test (%)*
			Correct	*Wrong*	*% correct*	*Correct*	*Wrong*	*% correct*	
Case	(a)	1140	162	105	61	610	263	70	65
	(b)	205	10	21	32	122	52	70	51
	(c)	1344	166	132	56	752	294	72	64

The overall fit of the model of the three samples is 65%, 51%, and 64% for (a), (b) and (c) respectively, indicating that the last two years were poorly forecast with only 32% correctly forecast unstable days. However the forecast for stable days was 70% correct which is reasonable. In comparison, the forecast for unstable days from Mt. Fidelity using only meteorological data (Section 8.7.3) is much more accurate. This could be because the dataset consisted of 30% unstable days in the Mt. Fidelity dataset and the fraction is thus much higher than the 22% in the Blue River dataset. It could also be that the different years for the test sample led to the poor results indicating that the winter of 2001/2002 may have been different from the other winters.

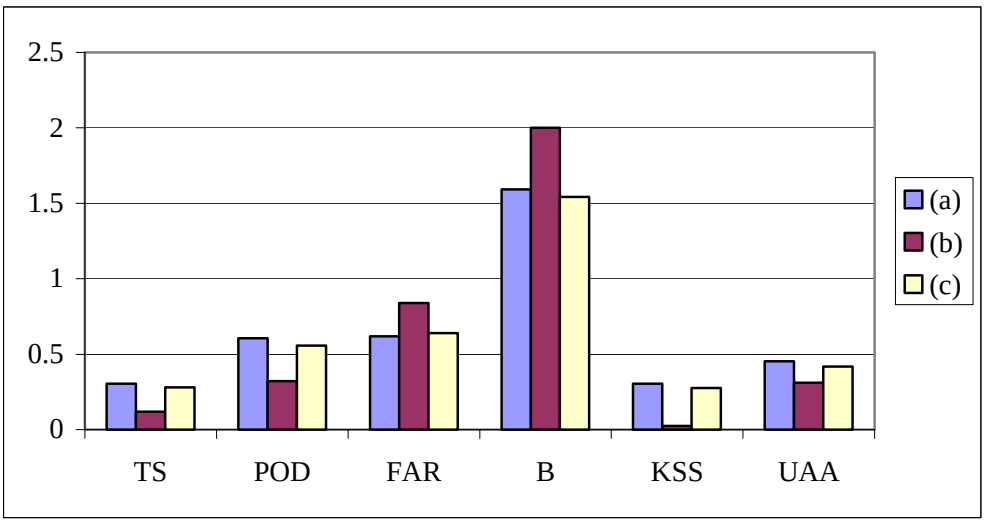

Figure 8.9: Forecast verification. Nearest neighbour analysis. Blue River.

The results of the forecast verification in Figure 8.9 indicates that the Blue River dataset was poorly forecast in the test sample with low scores for the threat score (TS), the probability of detection (POD) and the unweighted average accuracy (UAA). This indicates that the forecast for unstable days was unsatisfactory as a value of one represents a perfect forecast of unstable days. However the results for the learning dataset are much better, which could be a result of using the learning sample to build the model but could also reflect the similarity of the winters in that dataset. The bias (B) shows that unstable days were overforecast in all three samples.

As already observed in the Mt. Fidelity dataset, the forecast of skier-triggered avalanches on non-persistent weak layers using statistical models is difficult because skiing is often limited and releases are relatively small, leading to inconsistencies in the recordings. It also suggests that the definition of DSI is not as straightforward as for persistent weak layers.

8.9 Summary

To build on previous success with forecasting skier-triggered avalanches on persistent weak layers, various methods were applied to forecast skier-triggered avalanches on non-persistent weak layers. Shear frame testing was consistently done at Mt. Fidelity during storms and soon after. The data were used to build up two different datasets, one only including days during and soon after a storm when shear frame measurements were available and one daily dataset. A second study area was used to assess the performance of Cornice using only meteorological data.

It was not possible to define an empirical model to forecast the shear strength of storm snow layers over time and consequently, in order to calculate the stability index Sk38, shear frame measurements were necessary. However, none of the statistical tests proved that the skier stability index Sk38 improved the forecast of skier-triggered dry slab avalanches on storm snow instabilities in both dataset configurations. Thus it seems that there is no apparent value in measuring the shear strength with a shear frame on non-persistent weak layers to obtain a more accurate prediction of skier-triggered avalanches. However, the simple approach to defining Sk38 on stable days used in this analysis indicates some predictive value and it is likely to improve when consistent measurements during storms become available.

The Spearman rank correlations showed that most of the predictor variables correlated with the response variable (DSI), whereas multivariate effects were found only for the precipitation variables, the storm snow accumulation over the past three days and triggered avalanches in the classification tree analysis. However, more complex relationships with meteorological parameters are likely to exist.

Precipitation proved to be the most important predictor variable in the importance

rankings. This seems logical, because precipitation is directly related to the load and thus the quasi-static stress applied to the weak layers.

The Nearest Neighbour Models show the best results using meteorological data and the skier-triggered avalanches on the previous day, correctly forecasting 67%-70% of unstable days with skier-triggered avalanches on storm snow instabilities using Sk38 as a predictor, and 70%-74% of unstable days without Sk38, showing that unstable days are more accurately forecast.

The results of the configuration of Cornice for the second study area around Blue River does not show the same good results for forecasting unstable days as the first dataset. Unstable days are only forecast correctly in 32%-61% of the cases. However the percentage of unstable days in the dataset are higher for the first study area and better results are expected due to that fact. The overall performance of the models was 66%-68% for the Mt. Fidelity configuration and 51%-65% for the area around Blue River.

9. Forecasting natural avalanches on storm snow instabilities

9. 1 Introduction

Even though the focus of this thesis is to forecast *skier-triggered* avalanches, the value of a shear frame stability index to forecast *natural* avalanches (Sn38) will be assessed. The results of stability indices have been used at Rogers Pass for several years and are one factor for determining the stability of the snowpack. Based on this information, in addition to other factors, decisions are made on whether control shoots or road closures are necessary.

The aim of this analysis is to assess the importance of shear frame measurements for forecasting natural avalanches on storm snow instabilities on a daily basis. Storm snow instabilities may occur on persistent or non-persistent crystal forms.

9. 2 Dataset

The dataset is from Glacier National Park where shear frame measurements and meteorological data were available from Mt. Fidelity and Round Hill (see Figure 3.1, Chapter 3). The dataset consists of observations from 17 years (1987/1988 – 2003/2004) each starting at the first of December and ending 31st of March, except in 1990/1991, where consistent measurements were only available until the 8th of February. This time period was chosen because before the beginning of December shear frame measurements were rare, and in spring (April), avalanche activity differs from winter conditions. Altogether 2011 forecast days were considered. However not all predictor and response variables were available for each day, thus reducing the dataset accordingly.

9. 3 Methods

9.3.1 Response variables

Contrary to the analysis for skier triggering, a definition of an instability index is not necessary, because the avalanche observations are more consistent and do not depend on the amount of skiing. However, bad visibility in the forecast area affects the recording of

the avalanche activity, though the avalanche control section estimates the time of release on the next good visibility day.

The following response variables were chosen:

1) **Avalanche day (yes/no):** Avalanches of Size 1 and Size 2 are unlikely to threaten the highway through Glacier National Park and hence an avalanche day has to be defined accordingly. However a Size 2 avalanche might be a concern in some paths and consequently two definitions of an avalanche day are analysed:

AvalDay(1): A day on which one or more avalanches were reported in the forecast area of Size 2 or larger

AvalDay(2): A day on which one or more avalanches were reported in the forecast area of Size 3 or larger

Even though seventeen years of data were available for this study the avalanche control section at Rogers Pass only adopted the Canadian Avalanche Classification Scheme in the winter of 1995/1996. Thus avalanche days and maximum sizes of avalanches could only be defined for the last nine winters using the definitions above, reducing the datasets to nine years (1092 forecast days).

Table 9.1. gives on overview of the response variables.

Table 9.1. Summary of AvalDay(1) and AvalDay(2)

	Avalanche days	Non-avalanche days	N
AvalDay(1)	485	607	1092
AvalDay(2)	129	963	1092

2) **Maximum size of an avalanche**: The maximum size of an avalanche is the largest sized avalanche observed in the forecast area on a forecast day. Half size avalanches were grouped into the next lowest class, e.g. a Size 2.5 was classified as a Size 2 avalanche (Table 9.2).

Table 9.2. Summary of avalanche size class			
Avalanche Size	N	Size Class	N
0, 0.5	509	0	509
1, 1.5	98	1	454
2, 2.5	356	2	
3, 3.5	123	3	129
4, 4.5	6	4	
5	0	5	

Classification tree analysis using six possible classes would be rather complex and not necessary for avalanche forecasting purposes. The greatest interest is to determine whether the avalanche has the potential of reaching the highway or not. For this purpose it seems to be appropriate to classify the sizes into Size 0 – no avalanche occurrence, Size 1 and Size 2 – not likely to reach the highway and Sizes 3, 4 and 5 – potentially reaching the highway (Table 9.2). The possible advantage of three classes rather than two as in the analysis of avalanche days is that stable days differ in importance as regards smaller sized avalanches in evaluating the stability of the snowpack. The distribution of the number of avalanches recorded for each size reflects the recording standards at Rogers Pass. Size 1 avalanches are unlikely to leave the start zone and are not recorded during times of bad visibility and are also not as significant because they do not affect the highway. This explains why only 98 avalanches were reported over nine years.

3) **Avalanche index:** The avalanche index was introduced by Schweizer et al. (1998) and represents the sum of the mass of all observed avalanche occurrences. Accordingly, the various sizes, including half, sizes were given weights as in Table 9.3:

Table 9.3. Weighting scheme.	
Size	Weight
1	0.01
1.5	0.055
2	0.1
2.5	0.55
3	1
3.5	5.5
4	10
4.5	55
5	100

This weighting scheme implies that the mass increases 10 times from one size to another (CAA, 2002). After weighting all the avalanches on each day the days were classified into three classes with which to be able to run a classification tree analysis:

Table 9. 4. Classification of avalanche index values		
Class	Index Value	N
0	0	508
1	0 to 1	319
2	>= 1	265

Even though on some days the index was calculated with a value of up to 137 these three classes seem to be suitable because one Size 3 avalanche has a computed index of one and therefore is potentially dangerous to the highway. The number of observation days for all response variables was 1092 days.

The three response variables are similar, but combine different information. For example the avalanche index combines the number and the size of avalanches where 10 Size 2 avalanches are classified into the same class as 1 Size 3 avalanche. The avalanche day and the maximum size of an avalanche are based on the same premise that avalanches of a certain size may threaten the highway. However, while the avalanche day only considers two options, the maximum size distinguishes between three classes. Nearest

neighbour models only predict two classes and consequently only these two datasets for avalanche days can be used in the latter analysis.

9. 3.2 Predictor variables

The predictor variables used in this analysis are summarized in Table 9.5 and consist of variables measured at Mt. Fidelity or Round Hill. A.m. indicates that the measurements were collected in the morning, usually at 7 a.m., which is the time the forecasters generally retrieve their information.

Table 9.5. Daily predictor variables.

Abb.	**Definition**	**Fidelity**	**Round Hill**
T	Air temperature at a.m. (°C)	✓	~
WD	Wind direction at a.m.; (minus 90° modulo 360) east as base azimuth (°)	~	✓
RH	Relative humidity at a.m. (%)	✓	~
WS	Wind speed at a.m. (km h^{-1})	~	✓
WrunY	24 h wind run for previous day (km)	~	✓
HNY	Height of new snowfall for previous 24 h (m)	✓	~
HS	Height of snowpack at a.m. (m)	✓	~
PcpY	Water equivalent of precipitation on previous day (mm)	✓	~
Strm	Cumulative new snowfall (storm) since last day with less than 0.3 mm of precipitation (m)	✓	~
Strm3day	Cumulative new snowfall (storm) over the last three days (m)	✓	~
TminY	Minimum temperature for previous day (°C)	✓	~
TmaxY	Maximum temperature for previous day (°C)	✓	~
RHmnY	Minimum relative humidity for previous day (%)	✓	~
RHmxY	Maximum relative humidity for previous day (%)	✓	~
WSavg	Average wind speed over 24 hours	~	✓
WDavg	Average wind direction over 24 hours (minus 90° modulo 360) east as base azimuth	~	✓
Sn38	Stability index	✓	~
NaPrev	Number of natural avalanches on the previous day	Glacier National Park	

The availability of the predictor variables restricts the datasets in some of the analyses.

In order to calculate Sn38, shear strength measurements are required. The avalanche control section in Glacier National Park measures the shear strength in a level study plot at Mt. Fidelity mostly during and shortly after storms and discontinues once shear failures are not expected anymore or explosive avalanche control has been carried out. However, in daily avalanche forecasting models, daily shear strength values are necessary. In this study a simple approach is used where Sn38 is assigned a value of 10 (stable) when shear frame tests were attempted but no failures occurred, and also when the only failures were observed within 30 cm from the surface. Failures where the slab thickness was shallower than 30 cm are not considered as important in the stability evaluation of the avalanche control section (McMahon, personal communication, 2004). On days without measurements, Sn38 was also assigned a value of 10, because in most cases shear frame tests were not done because failures were unlikely. However, a false stable error may occur, because on some days, for example during control shoots, when time was a constraint, testing may have been omitted. On 528 days (48% of the 1092 days) shear strength measurements were available and not assigned a value of 10, with a minimum of 0.117 kPa, a maximum of 8.627 kPa, a mean of 1.374 kPa and a standard deviation of 0.77 kPa.

It is clear that assigning a value of 10 to days where failures were unlikely is a simple approach and to formulate a forecasting model requires more research. However an attempt was made to assess the value of shear frame measurements to predict natural dry slab avalanches in a daily forecasting model.

9.4 Avalanche day (yes/no)

Rank correlations between the predictor variables and AvalDay(1) and AvalDay(2) are listed in Tables 9.6 and 9.7.

Table 9.6. Spearman rank correlations with AvalDay(1)			
Variable	**Valid N**	**Spearman R**	**p-level**
PcpY	1057	0.390	$<1E10^{-17}$
HNY	1084	0.344	$<1E10^{-17}$
RH	1083	0.310	$<1E10^{-17}$
T	1082	0.293	$<1E10^{-17}$
NaPrev	1083	0.288	$<1E10^{-17}$
Strm3day	1059	0.283	$<1E10^{-17}$
RhminY	1084	0.267	$<1E10^{-17}$
Sn38+	1092	-0.261	$<1E10^{-17}$
Strm	1064	0.255	$3.0E10^{-17}$
RhmaxY	1084	0.250	$6.0E10^{-17}$
TminY	1084	0.242	$7.3E10^{-16}$
WDavg	1092	-0.182	1.4E-9
WS	1092	0.175	5.4E-9
Wsavg	1092	0.162	7.8E-8
WrunY	1092	0.162	7.8E-8
TmaxY	1084	0.150	6.6E-7
HS	1092	0.103	6.4E-4
WD	1085	-0.089	0.003

Table 9.7. Spearman rank correlations with AvalDay(2)			
Variable	**Valid N**	**Spearman R**	**p-level**
PcpY	1057	0.341	$<1E10^{-17}$
HNY	1084	0.307	$<1E10^{-17}$
Strm3day	1059	0.299	$<1E10^{-17}$
NaPrev	1083	0.258	$1.0E10^{-17}$
Strm	1064	0.257	$2.0E10^{-17}$
T	1082	0.246	$2.1E10^{-16}$
RhminY	1084	0.236	$4.0E10^{-15}$
RH	1083	0.224	$7.9E10^{-14}$
Sn38+	1092	-0.208	$3.5E10^{-12}$
TminY	1084	0.182	$1.5E10^{-9}$
WS	1092	0.164	5.1E-8
RhmaxY	1084	0.162	$8.9E10^{-8}$
TmaxY	1084	0.136	6.6E-6
Wsavg	1092	0.124	4.2E-5
WrunY	1092	0.123	4.5E-5
WDavg	1092	-0.122	5.5E-5
HS	1092	0.115	1.5E-4
WD	1085	-0.104	5.7E-4

All of the 18 predictor variables were significantly correlated at a 5% significance level in both datasets, however the significance of each predictor variable on the response

variable in the second dataset (AvalDay(2)) decreased. The two highest ranked predictors in both datasets are the precipitation and the new snow height over the past 24 hours. The height of the snowpack and the wind direction in the morning only exhibit a weak correlation. All variables, except Sn38, WDavg and WD correlated positively with the response variables in both datasets.

Positive correlations

Precipition (PcpY), height of new snow (HN), storm snow (strm) and storm snow accumulations over the past 3 days (strm3day): The positive correlation of these predictor variables imply that avalanches of Size 2 or larger are more likely to release when the amount of precipitation increases. However the two storm snow predictors (strm and strm3day) are more strongly correlated with AvalDay(2) while PcpY and HN are more strongly correlated with AvalDay(1).

Relative humidity (RH, RHmin, RHmax): All three humidity variables show a positive correlation with AvalDay(1) and AvalDay(2). However the correlations are strongest with AvalDay(1) while RH has the third strongest correlation. Higher humidity is generally observed during times of precipitation, and during stormy periods avalanches are more likely. The humidity measured in the morning was most highly correlated in the AvalDay(1) dataset, whereas RHmin was more strongly correlated in the AvalDay(2) dataset.

Temperature (T, Tmin and Tmax): The air temperature is positively correlated with the response variable in both datasets implying that during warmer periods more avalanches are likely. During storms the temperature is generally warmer than during clear sky conditions between storms in the winter.

Wind speed and wind run (WS, WSavg, Wrun): The ridge top wind speed and the wind run are all positively correlated with the response variables. During times with higher wind speeds more avalanches are likely because of the wind transportation, the developments of wind slabs and greater loading of start zones.

Number of avalanches on the previous day (NaPrev): The number of natural avalanches on the previous day is strongly correlated in both datasets implying that the

more avalanches that occurred on the previous day, the more likely it is that an avalanche will occur on the forecast day. This is possibly because storms often last longer than one day and the slab becomes more cohesive over time.

Negative correlations

Wind direction (WDavg, WD): The wind direction is measured on a 360° scale with 0° indicating north, 90° east, 180° south and 270° west. The negative correlation of the average wind direction over the past 24 hours and the wind direction measured in the morning implies that an avalanche day is more likely when the wind direction has a lower angle from North. This is consistent with the wind shift during warm fronts in the northern hemisphere.

Sn38: The stability index for natural dry slab avalanches, Sn38, correlated negatively, indicating that at lower values more avalanches are likely, as can be expected. However, in the first dataset the correlation is stronger, which means that Sn38 is less predictive for larger avalanches.

The entire dataset was divided into a learning sample consisting of data from the winters 1995/1996 to 2001/2002 and a test sample, which comprises data from the past two winters (2002/2003 and 2003/2004). Table 9.8 summarizes the sample sets as used in the development of the classification tree after 52 days were deleted due to missing values of some predictor variables.

Table 9.8. Summary of the test and learning sample in classification tree.

	AvalDay(1)		AvalDay(2)	
	Avalanche days	Non-avalanche days	Avalanche days	Non-avalanche days
Learning sample (N = 814)	354	460	94	720
Test sample (N = 220)	109	111	29	191

As explained previously, by putting a higher prior on classes, which are more important for forecasting correctly, the development of the tree can be influenced accordingly. Similar to the analysis of skier-triggered avalanches, it is more important to forecast avalanche days than stable days. However there is no easy way to find the priors

that are most suitable except by trial and error. The aim is to set the priors so that misclassified avalanche days are at a minimum, but that the number of misclassified non-avalanche days are not unreasonably high (below 60%). The best classification in the learning sample was achieved by setting the priors to 0.4 for non-avalanche days and 0.6 for avalanche days.

Of 18 possible predictor variables, only two splitting variables in three splits were used in the classification tree in Figure 9.1.

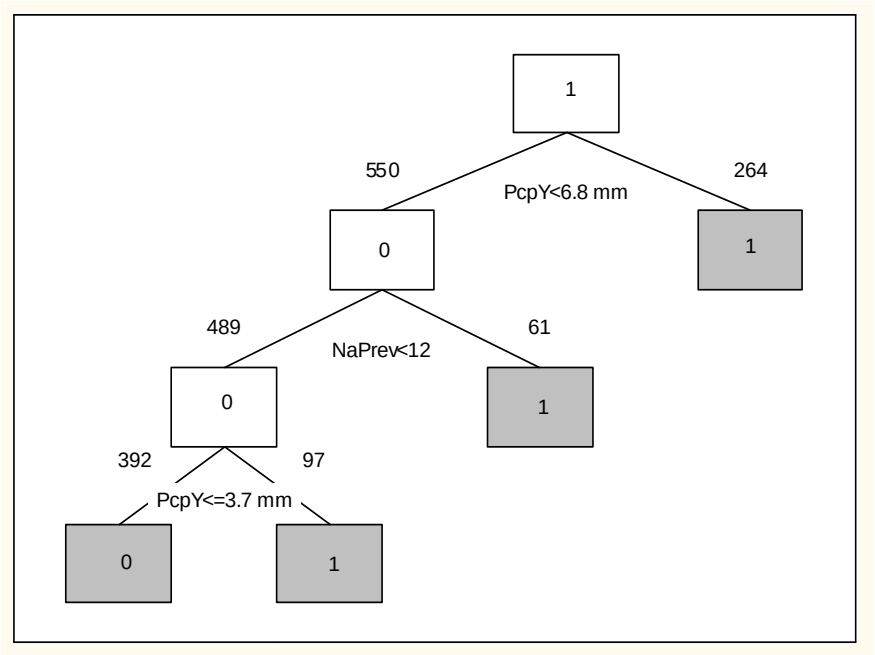

Figure 9.1: Classification Tree, AvalDay(1).

In the first split, avalanche days are classified when the precipitation over the past 24 hours is equal to or greater than 6.8 mm (right branch). On days with less than 6.8 mm (left branch) the next split determines avalanches when more than 12 natural avalanches were reported on the previous day (right branch, second split). In the case of fewer than 12 observed avalanches on the previous day, one last split classifies avalanche days when the precipitation over the past 24 hours was equal to or exceeded 3.7 mm (right branch, third

split). Less than 3.7 mm precipitation on a forecast day indicates non-avalanche days (left branch, third split). Including and excluding Sn38 resulted in an identical classification tree.

The global cross-validation results of the learning and the test sample are shown in Table 9.9.

Table 9.9. Global cross-validation results for the learning sample and misclassification results for test sample. AvalDay(1)

	Misclassification			
	Avalanche days	Non-avalanche days	Cost	Error
Learning sample	115/354 (33%)	143/460 (31%)	0.319	0.017
Test sample	43/109 (40%)	35/111 (32%)	0.363	0.033

The model fits 67% of the avalanche days and 69% of the non-avalanche days in the learning sample. In the test sample avalanche days were forecast less accurately (60%) and non-avalanche days with 68%, only 1% less accurately than in the learning sample. The misclassification cost in the test sample was with 0.363, higher than in the learning sample, which had a value of 0.319.

The predictor variables were ranked on a scale from 0 to 100 according to their importance in the classification tree development. The precipitation over the past 24 hours was ranked highest, which was expected and corresponds with the result from the Spearman rank correlation. HN correlated strongly with the response variable in the Spearman rank correlation analysis, although in the multivariate approach it was only ranked eleventh out of 18 variables considered. Sn38 ranked seventh and did not contribute to the classification of AvalDay(1).

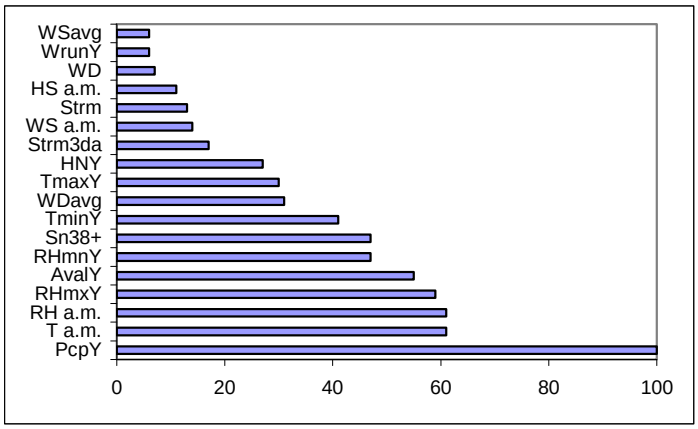

Figure 9.2: Importance ranking. AvalDay(1).

In the second tree (Figure 9.3), the classification of the response variable AvalDay(2) was made. Only one split led to two terminal nodes where an avalanche day was predicted in the case of 8.7 mm of precipitation or more over the past 24 hours. None of the other 18 remaining predictor variables was used for the development of the tree. Again this underlines the predominance of the precipitation, especially for larger sized avalanches. The priors were set to 0.35 and 0.65.

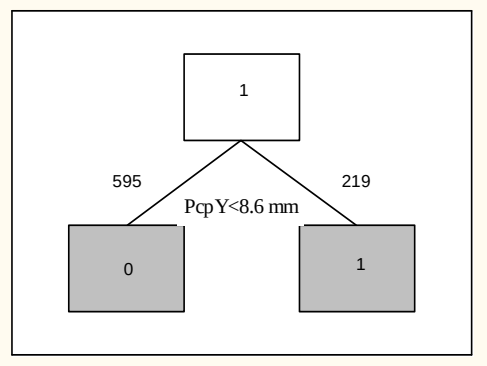

Figure 9.3: Classification tree. AvalDay(2).

The tabulated results form the global cross-validation in Table 9.10 are more promising than the results from the first classification tree.

Table 9.10. Global cross-validation results for the learning sample and misclassification results for test sample. AvalDay(2).

	Misclassification			
	Avalanche days	Non-avalanche days	Cost	Error
Learning sample	25/94 (27%)	237/720 (33%)	0.288	0.03
Test sample	9/29 (31%)	30/191 (16%)	0.257	0.057

The classification tree predicted 73% of the avalanche days in the learning sample and 69% in the test sample correctly. The correct classification for the non-avalanche days was computed with 67% and 84% in the learning and the test sample respectively. The misclassification cost in the learning sample equalled 0.288 and 0.257 in the test sample. However the same result was achieved when excluding Sn38 and again the predictive value of Sn38 is this multivariate model is not apparent.

The importance ranking in Figure 9.4 again shows precipitation as the most important predictor, even more dominant than in the tree for AvalDay(1). Only four of the predictors on a scale of 0 to 100 have a value of more than 30, indicating that the relative importance of most of the predictors is low in relation to three precipitation variables and temperature.

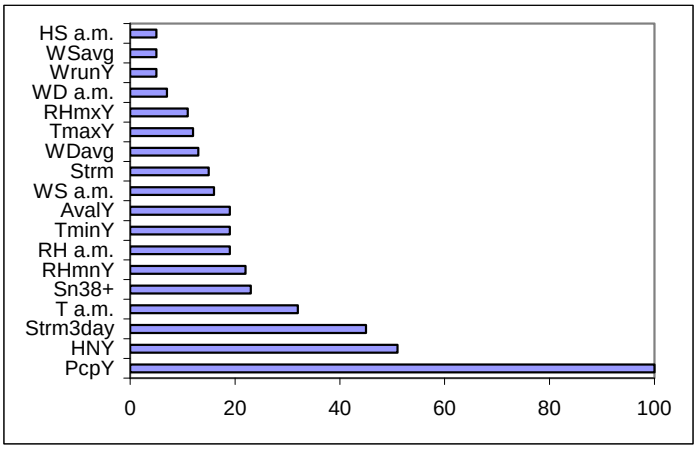

Figure 9.4: Importance ranking. AvalDay(2).

The classification tree results showed that avalanche days with Size 3 and larger avalanches (AvalDay(2)) can be predicted reasonably using the tree model. However the classification of Size 2 and larger avalanches (AvalDay(1)) was not as good. Also Sn38 did not contribute to the fit of either tree model, though a similar shear frame stability index is considered to be a good predictor in operational forecasting at Glacier National Park. This suggests that simple discrimination into avalanche and non-avalanche days may not be the right approach. It was also shown that days with larger avalanches can be better predicted and days with smaller avalanches have different predictor importance. This leads to the definition of the next response variable, which is based on the maximum size of an avalanche occurrence on a given day.

9.5 Maximum size class (MaxSz)

As in the analysis for avalanche days, 52 observations had to be deleted because of missing predictor variables. The sample sets of MaxSz is summarized in Table 9.11.

Table 9.11. Summary of the learning and test sample. AvalMax.

	Class 0	Class 1	Class 2	N
Learning sample	378	342	94	814
Test sample	98	93	29	220

The Spearman rank correlations (Table 9.12) are similar to the ones for avalanche days with the same variables having positive and negative correlations, as expected. However a stronger correlation was calculated for strm3day and Sn38, indicating that the more snow that fell in the past three days and the lower the value of Sn38, a larger maximum avalanche could be expected.

Table 9.12. Spearman rank correlations with maximum size class of avalanche.

Variable	Valid N	Spearman R	p-level
PcpY	1057	0.429	$<1E10^{-17}$
HNY	1084	0.388	$<1E10^{-17}$
RH	1083	0.321	$<1E10^{-17}$
Strm3day	1059	0.319	$<1E10^{-17}$
NaPrev	1083	0.314	$<1E10^{-17}$
Sn38+	1092	-0.311	$<1E10^{-17}$
T	1082	0.311	$<1E10^{-17}$
Strm	1064	0.295	$<1E10^{-17}$
RhminY	1084	0.292	$<1E10^{-17}$
RhmaxY	1084	0.275	$<1E10^{-17}$
TminY	1084	0.258	$1.0E10^{-17}$
WS	1092	0.212	1.5E-12
WDavg	1092	-0.180	2.2E-9
WSavg	1092	0.176	4.8E-9
WrunY	1092	0.176	4.9E-9
TmaxY	1084	0.171	1.5E-8
HS	1092	0.124	4.3E-5
WD	1085	-0.100	9.7E-4

Three out of 18 predictor variables were used in the development of the classification tree. These are the 24 hour precipitation, Sn38 and the temperature measured in the morning (Figure 9.5).

Deleted: ¶

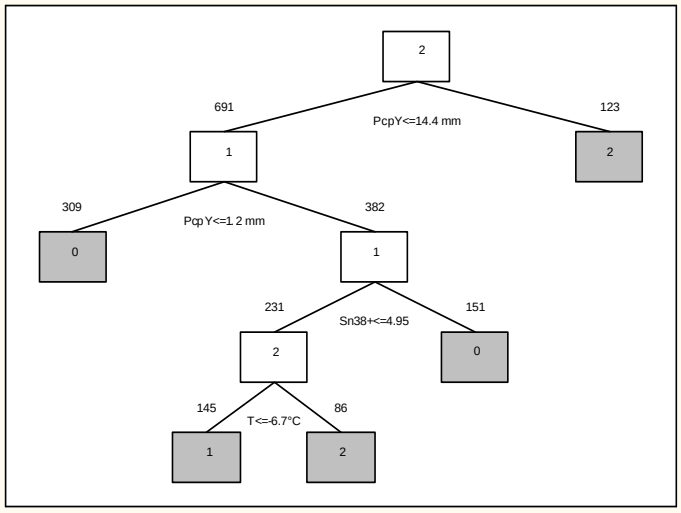

Figure 9.5: Classification tree to predict the maximum size of an avalanche

Similar to the classification of avalanche days, PcpY contributed most to the fit of the classification tree. In the first split Class 2 was predicted when the precipitation over the past 24 hours was greater than 14.38 mm (right branch). The remaining days (left branch) were classified as Class 0 when the measured precipitation did not exceed 1.23 mm within 24 hours (left branch, second split) and the right branch (second split) was again split and classified as Class 0 when Sn38 was greater or equal to 4.96 (right branch, third split). In the last split, Class 1 was predicted at temperatures colder than –6.72°C and Class 2 when the temperatures were warmer.

The three splitting variables in the classification tree are also ranked highest in importance ranking in Figure 9.6. The importance of Sn38 is ranked in third position, higher than in the Spearman rank correlation where it had only the sixth strongest correlation, indicating that the value of Sn38 increased in a multivariate approach.

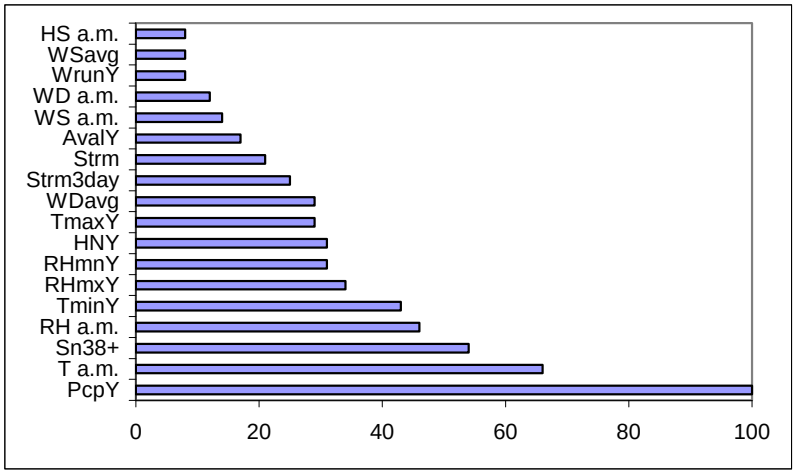

Figure 9.6:Predictor importance. MaxSz including Sn38.

The results of the global cross-validation for the learning and the test samples are summarized in Tables 9.13 and 9.14. Figures 9.7 and 9.8 show the predicted versus the observed classes. A lower predicted maximum size than that observed is considered false stable forecast whereas a higher predicted class is a conservative forecast.

In both the learning and the test sample the classification of Class 1 is poor with only 28% and 19% correctly forecast respectively. Class 1 was misclassified as Class 0 in 36% of the days in the learning sample and in 54% in the test sample, which are false stable forecasts. This poor classification of Class 1 may be because avalanches of Size 1 are less often reported than expected.

Class 2 was correct in 64% in the learning sample and in 62% in the test sample. The misclassification of Class 2 as Class 1 was 20% and 3% for the learning and test samples, respectively. Sixteen per cent and 28% of the observations were classified as Class 0, which could have adverse consequences.

Correct prediction of Class 0 was 63% in the learning and 84% in the test samples, a surprisingly good result considering that the priors were set to 0.28, 0.32 and 0.4 for Class 0, Class 1 and Class 2 respectively. Different settings did not improve the fit. The estimated misclassification cost equalled 0.481 in the learning sample and 0.456 in the test sample.

Table 9.13. Global cross-validation results for the learning sample MaxSz with Sn38

	Misclassification (Cost 0.481, Error 0.022)		
	0	1	2
0	237 (63%)	122 (36%)	15 (16%)
1	80 (21%)	94 (28%)	19 (20%)
2	61 (16%)	126 (37%)	60 (64%)
Total	378	342	94

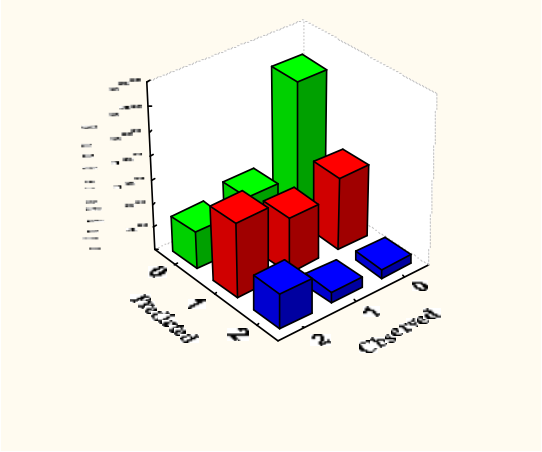

Figure 9.7: Predicted versus observed class for learning sample.

Table 9.14. Global cross-validation results for the test sample. MaxSz with Sn38			
	Misclassification (Cost 0.456, Error 0.04)		
	0	1	2
0	82 (84%)	50 (54%)	8 (28%)
1	7 (7%)	18 (19%)	3 (10%)
2	9 (9%)	25 (27%)	18 (62%)
Total	98	93	29

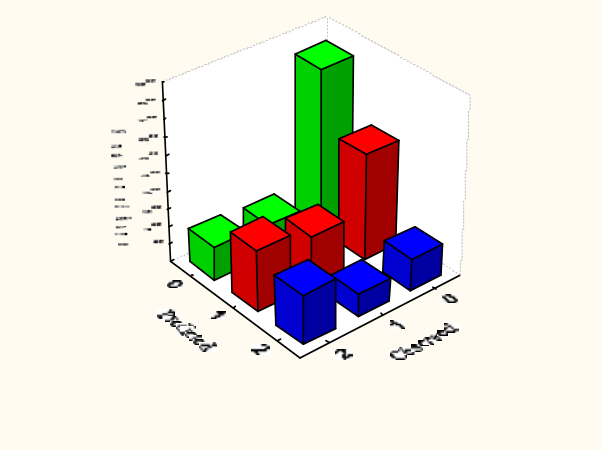

Figure 9.8: Predicted versus observed class for test sample.

Excluding Sn38 improved the overall fit of the model slightly (Tables 9.15 and 9.16) with a lower cost of 0.464 compared to 0.481 in the learning sample and 0.433 compared to 0.456 in the test sample. However, the predictions for Class 2 were worse, with 61% correct classifications in the learning compared to 64%, and 48% in the test sample compared to 62% when including Sn38. The better cost is because of the improved forecast for Class 1, with 39% correct forecast compared to 28% in the learning sample, and 62% compared to 19% in the test sample. The forecast for non-avalanche days was worse in the learning sample with 61% compared to 63%, and in the test sample with 62% compared to 84%. However the classification tree using Sn38 is preferred, because the larger sized avalanches are better predicted, even though the contribution of Sn38 is only small.

Table 9.15. Global cross-validation results for the learning sample					
	Misclassification				
	0	1	2	Cost	Error
0	229 (61%)	106 (31%)	10 (11%)	0.464	0.023
1	105 (28%)	132 (39%)	27 (29%)		
2	44 (12%)	104 (30%)	57 (61%)		
Total	378 (100%)	342 (100%)	94 (100%)		

Table 9.16. Global cross-validation results for the test sample					
	Misclassification				
	0	1	2	Cost	Error
0	61 (62%)	25 (27%)	3 (10%)	0.433	0.043
1	32 (33%)	58 (62%)	12 (41%)		
2	5 (5%)	10 (11%)	14 (48%)		
Total	98 (100%)	93 (100%)	29 (100%)		

The classification tree is the same as the one that includes Sn38, except that the tree only consists of the first two splits based on the precipitation over the past 24 hours.

The results suggest that Sn38 is a better predictor for larger sized avalanches and that Class 1, including Sizes 1 and 2 are not well forecast, possibly because Sn38 was assigned a value of 10 for slab depths shallower than 30 cm and Size 1 avalanches are not recorded consistently.

9.6 Avalanche index

As shown in Table 9.17, the Spearman rank correlations for the Avalanche Index are similar to the other analysed datasets, though Sn38 and strm3day are more strongly correlated with the avalanche index than with the maximum size and avalanche days.

Table 9.17. Spearman rank correlations with maximum size class of avalanche. AvalIndex.

Variable	Valid N	Spearman R	p-level
PcpY	1057	0.454	$<1E10^{-17}$
HNY	1084	0.409	$<1E10^{-17}$
Sn38+	1092	-0.332	$<1E10^{-17}$
Strm3day	1059	0.332	$<1E10^{-17}$
RH a.m.	1083	0.328	$<1E10^{-17}$
NaPrev	1083	0.318	$<1E10^{-17}$
T a.m.	1082	0.315	$<1E10^{-17}$
RhminY	1084	0.310	$<1E10^{-17}$
Strm	1064	0.301	$<1E10^{-17}$
RhmaxY	1084	0.282	$<1E10^{-17}$
TminY	1084	0.250	$6.0E10^{-17}$
WS(RH)	1092	0.224	7.45E-14
WDavg(RH)	1092	-0.205	9.0E-12
WrunY(RH)	1092	0.190	2.3E-10
WSavg(RH)	1092	0.190	2.3E-10
TmaxY	1084	0.152	4.7E-7
HS	1092	0.114	1.7E-4
WD(RH)	1085	-0.097	0.001

The classification tree for the Avalanche Index in Figure 9.9 used five splitting variables (PcpY, NaPrev, Strm, Sn38, Wdavg), in seven splits.

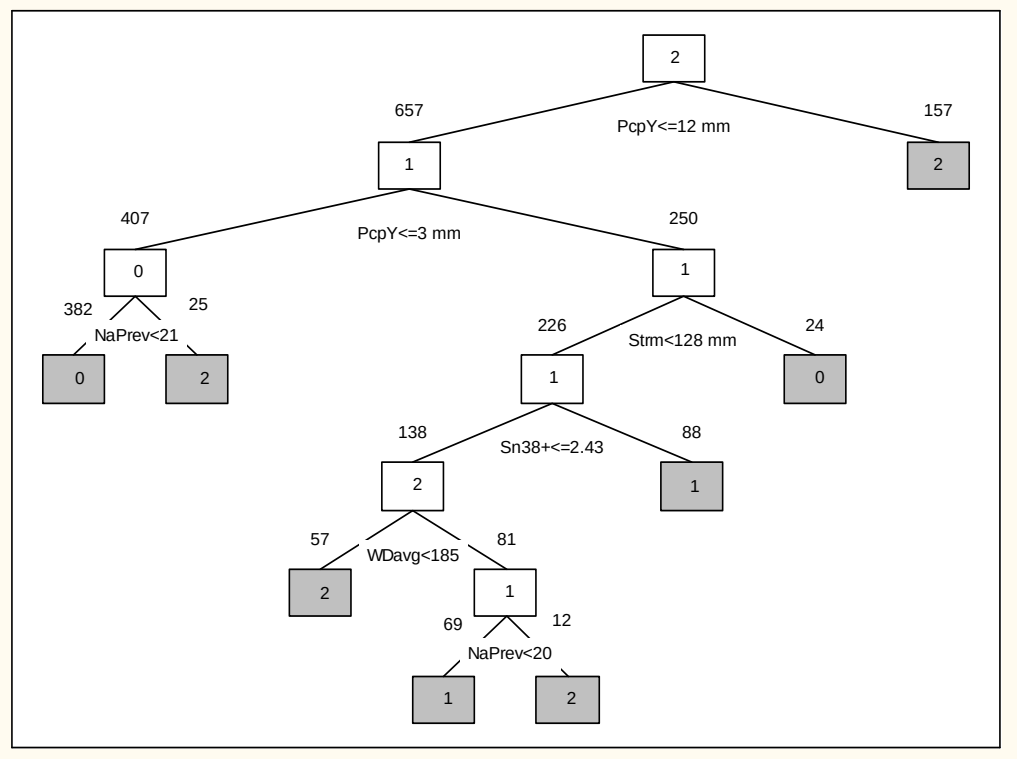

Figure 9.9: Classification tree. AvalIndex, including Sn38

Similar to all the other trees, the precipitation over the past 24 hours was used in the first split. However, the maximum size of the splitting value is 11.95 mm whereas it is 14.38 mm for the maximum size predicting Class 2 (right branch). In the second split, days with less than 11.95 mm of precipitation are again divided. This time precipitation values less than 3 mm (left branch, second split) lead into the third split, where Class 0 is classified when less than 22 avalanches were reported on the previous day (third split, left branch). However where more than 22 avalanches are observed, Class 2 is likely. The right branch of the second split leads to split four, where storm snow accumulations of 128 mm or more predict Class 0, which was not expected because the Spearman rank correlation analysis showed that more storm snow leads to a higher AvalIndex (split four, right branch). The left branch of split four goes into split five where Sn38 values of 2.43 or greater classify Class 1 avalanches (right branch) and smaller values of the stability index

indicate Class 2 where the wind direction was south-east, east or northerly (sixth split, left branch). However the right branch of the sixth split “predicts” Class 2 when more than 20 avalanches were observed on the previous day and Class 1 with fewer avalanches observed (split 7, left branch).

The importance ranking in Figure 9.10 reveals that precipitation variables were not as dominant when classifying the avalanche index compared to the classification of avalanche days and maximum sizes. Six additional variables (T, Sn38, RH, RHmax, WDavg and NaPrev in decreasing importance) had a relative importance of 50 or more on a scale of 0 to 100.

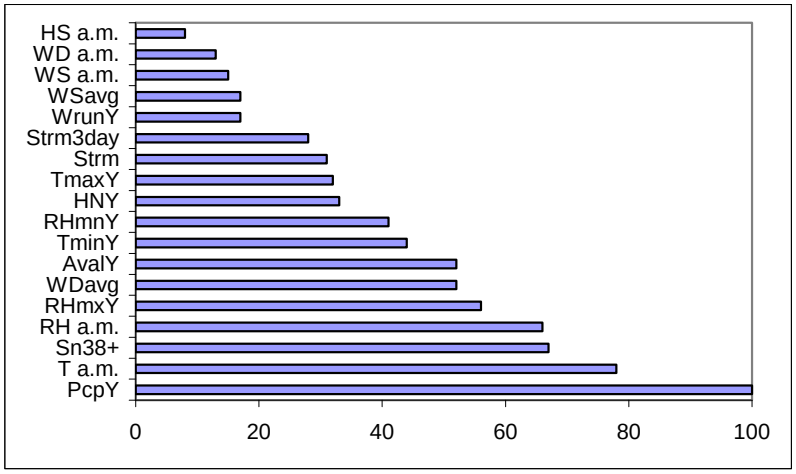

Figure 9.10: Importance ranking. AvalIndex.

The results from the global cross-validations are summarized in Tables 9.18 and 9.19 and Figures 9.11 and 9.12 for the learning and the test sample.

Table 9.18. Global cross-validation results for the learning sample AvalIndex including Sn38			
	Misclassification (cost 0.457, error 0.019)		
	0	1	2
0	231 (61%)	91 (39%)	32 (16%)
1	108 (29%)	104 (44%)	54 (27%)
2	39 (10%)	41 (17%)	114 (57%)
Total	378	236	200

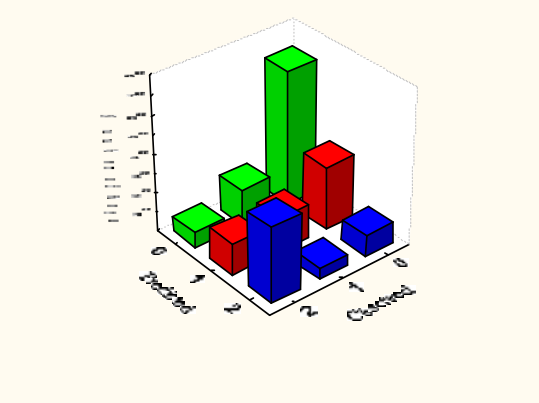

Figure 9.11: Predicted versus observed class for learning sample.

Table 9.19. Global cross-validation results for the test sample AvalIndex including Sn38			
	Misclassification (Cost 0.425, Error 0.033)		
	0	1	2
0	70 (72%)	33 (49%)	11 (20%)
1	19 (20%)	18 (27%)	6 (11%)
2	8 (8%)	16 (24%)	39 (70%)
Total	97	67	56

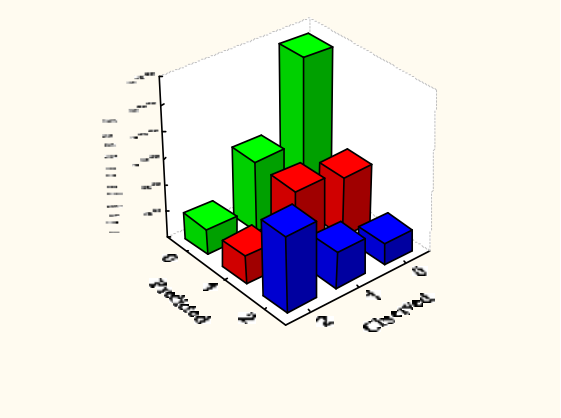

Figure 9.12: Predicted versus observed class for test sample.

The classification of the avalanche index is not ideal either. Again for Class 1 the misclassification was 56% and 73% in the learning and test sample respectively. Poor and false unstable days were classified in 39% and 49% of the observations. However, the test sample performs better than the learning sample to classify Class 2 with a rate of 70% and 57% respectively. The classification for Class 0 is acceptable with 61% correctness in the learning and 72% in the test sample.

The misclassification cost equalled 0.457 in the learning sample and 0.425 in the test sample. Excluding Sn38 from the analysis resulted in a higher cost with 0.47 in the learning sample, though producing a lower cost in the test sample with 0.382 (Table 9.20 and 9.21, Figures 9.14 and 9.15). In addition, Class 2 indices are better classified, which is not in accordance with the maximum size analysis where larger avalanches were predicted better when including Sn38.

Excluding Sn38 led to a similar tree structure in the upper part of the tree, which is most influential for the fit of the dataset (Figure 9.13).

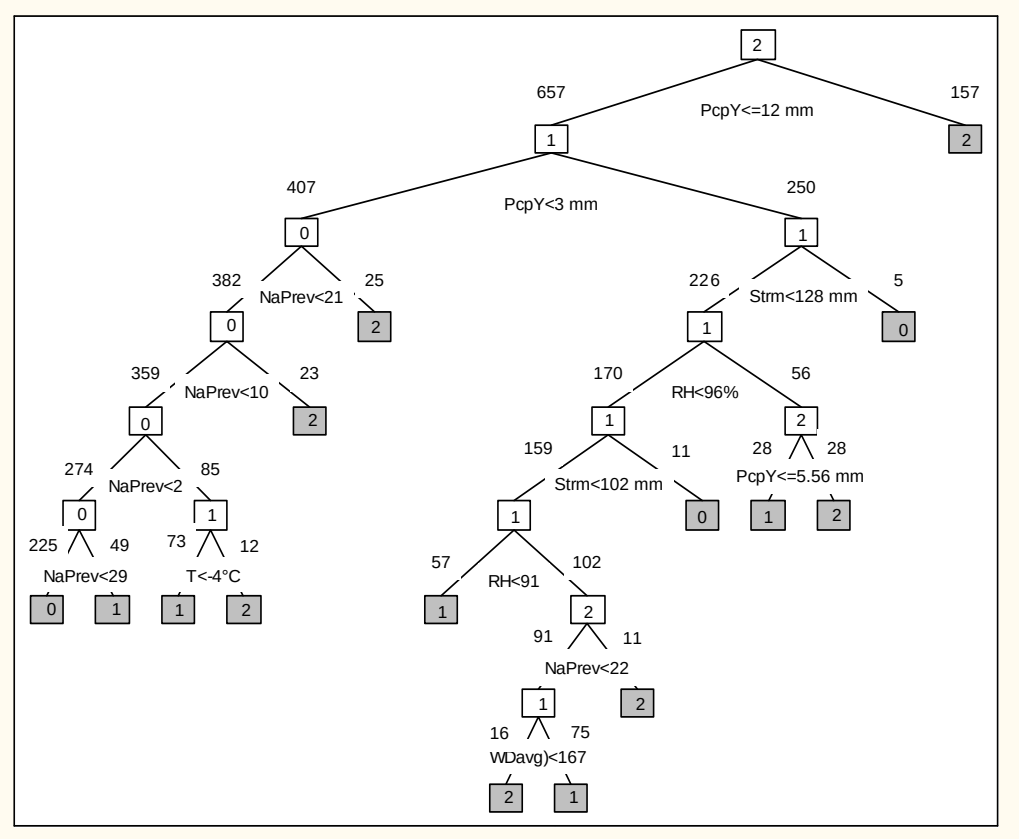

Figure 9.13: Classification tree for AvalIndex excluding Sn38.

Consequently the differences in the selected predictors were minor after the first splits. However, the complexity of the tree increased when Sn38 was excluded, resulting in 14 splits and six splitting variables (PcpY, NaPrev, Strm, RH, T and Wdavg). However the expected performance of both trees is similar because the first four splits are identical.

As already mentioned, the cost was higher in the learning sample when excluding Sn38 though lower in the test sample. Apart from the analysis of the maximum size avalanche Sn38 did not improve the classification of larger sized avalanches. The exclusion of Sn38 predicted Class 2 4% more accurately (61% compared to 57%) in the learning sample and 3% more accurately in the test sample (73% compared to 70%). Also, Class 1 was better predicted in the test sample (54% compared to 27%) and the false stable classifications for Class 1 were reduced in both the test and the learning sample to 28% and 2% respectively. However Class 0 was correctly classified by 60% in the learning sample and 54% in the test sample, which is reasonable considering the overall poor results.

Table 9.20. Global cross-validation results for the learning sample. AvalIndex, excluding Sn38

	Misclassification (Cost 0.47, error 0.019)		
	0	1	2
0	226 (60%)	88 (37%)	31 (16%)
1	86 (23%)	84 (36%)	47 (23%)
2	66 (17%)	64 (27%)	122 (61%)
Total	378	236	200

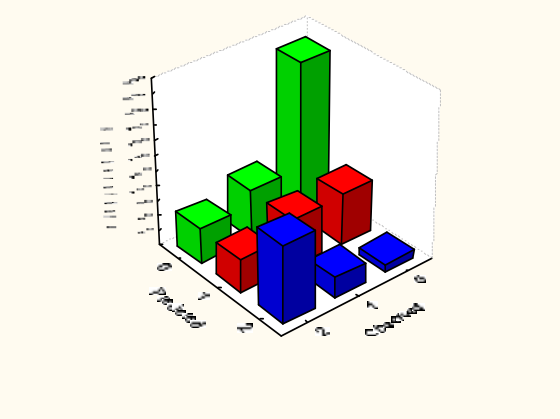

Figure 9.14: Predicted versus observed class for learning sample.

Table 9.21. Global cross-validation results for the test sample. AvalIndex, excluding Sn38			
	Misclassification (Cost 0.382, error 0.034)		
	0	1	2
0	52 (54%)	14 (21%)	3 (5%)
1	30 (31%)	36 (54%)	12 (22%)
2	15 (15%)	17 (25%)	41 (73%)
Total	97	67	56

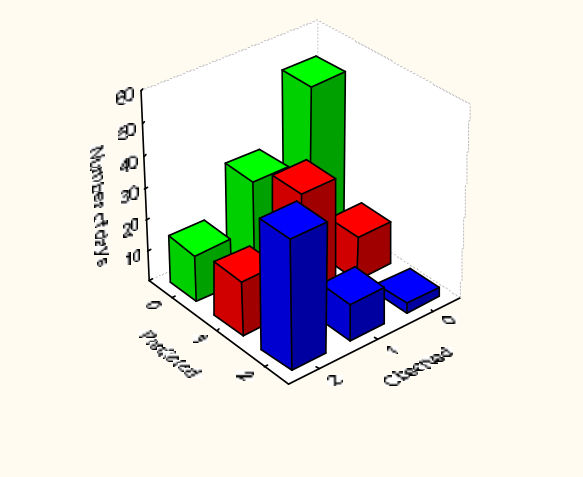

Figure 9.15: Predicted versus observed class for test sample.

As already observed the importance of precipitation is less when classifying the avalanche index, and the other variables gain more importance. This is possibly because the index is more complex in this definition.

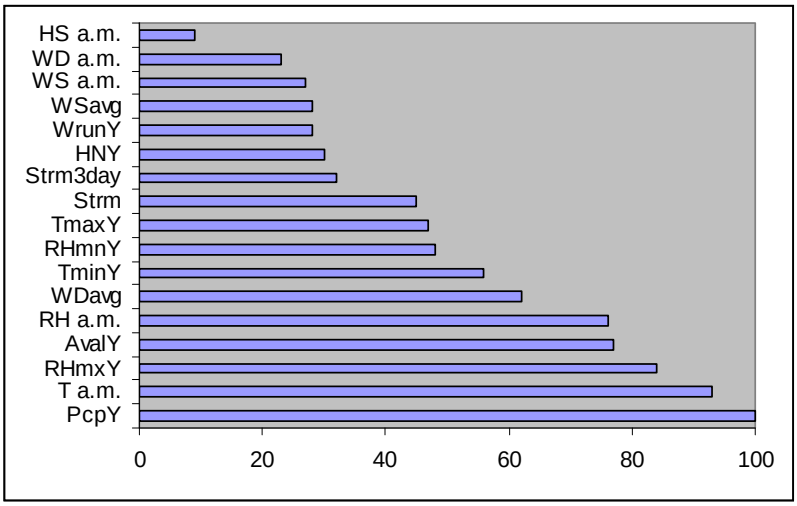

Figure 9.16: Importance ranking. AvalIndex, excluding Sn38

Although the misclassification is rather high for all the response variables, it must be remembered that avalanches are difficult to forecast, and the models presented are, at most, only intended to provide an independent forecast for consideration by an experienced forecaster. However Sn38 appeared to be important in some of the analysis, thought not as dominantly as for forecasting skier-triggered avalanches. This was only a simple approach to the problem and the results can certainly be improved. Daily shear frame measurements would be important in order to be able to formulate accurately a strength model.

Although the misclassification is rather high for all the response variables, it must be remembered that avalanches are difficult to forecast, and the models presented are, at most, only intended to provide an independent forecast for consideration by an experienced forecaster. However Sn38 appeared to be important in some of the analysis, thought not as dominantly as for forecasting skier-triggered avalanches. This was only a simple approach to the problem and the results can certainly be improved. Daily shear frame measurements would be important in order to be able to accurately formulate a strength model.

9.7 Nearest Neighbour model

The Nearest Neighbour model was configured separately for the response variables AvalDay(1) and (2). All other response variables have more than two classes and thus cannot be used in Cornice. As in the preceding chapters (a) stands for the learning, (b) for the test and (c) for the combined samples. Table 9.22 shows the results for AvalDay(1). Cornice was configured with the learning sample using the automatic weighting function. However the learning sample is more balanced with 44% avalanche days than the learning samples used for the analysis of skier-triggered avalanches. Cornice uses the three nearest neighbours as the threshold for a correct forecast when optimizing the weights and consequently weighting for correctly forecasting avalanche days over non-avalanche days and also accounting for unbalanced datasets. However this results in an overforecast of avalanche days in Cornice when the dataset is more or less balanced. Unfortunately it is not possible to change the weighting scheme to optimize the forecast with more or less than three nearest neighbours. However it is possible to calculate the fit using a different number of nearest neighbours by analysing the log file, though the weights are optimized to forecast an avalanche day when three or more nearest neighbours are avalanche days. Table 9.22 shows how the number of nearest neighbours chosen influences the outcome of the forecast due to the balance of a dataset.

Table 9.22. Performance of Nearest Neighbour model. AvalDay(1).

		Days forecast	Avalanche days (DSI = 1)			Non-avalanche days (DSI = 0)			Fitness of batch test (%)
			Correct	Wrong	% correct	Correct	Wrong	% correct	
With	(a)	821	320	37	90	171	293	37	63
Sn38 3NN	(b)	237	104	14	88	40	79	37	63
	(c)	1047	430	45	91	179	405	31	61
Without	(a)	821	317	40	89	160	304	35	62
Sn38 3NN	(b)	237	104	14	88	51	68	43	66
	(c)	1065	424	52	89	198	391	34	61
With	(a)	821	229	128	64	291	173	63	63
Sn38 5NN	(b)	237	63	55	53	85	34	71	62
	(c)	1047	300	175	63	366	218	63	63
Without	(a)	821	213	144	60	314	150	68	64
Sn38 5NN	(b)	237	65	53	55	91	28	77	66
	(c)	1065	274	202	58	409	180	69	64

The percentage of correctly avalanche days is 90%, 88% and 91% for (a), (b) and (c) respectively, very high when including Sn38 and also when excluding Sn38 with 89%, 88% and 89% fit for avalanche days. However the correct forecast for non-avalanche days in all three cases yielded a fit of only 31%-43%. This reflects the unsuitability of a balanced dataset in Cornice using the default of three nearest neighbours. A threshold of five nearest neighbours was considered reasonable because 40% of the dataset consists of avalanche days and using five or more accounts for the difference in the dataset and also for the importance of forecasting avalanche days more correctly. The results using five nearest neighbours as the threshold for forecasting an avalanche day are much more realistic and are comparable to the classification tree results. However the results are less accurate in the learning sample with 36% misclassified avalanche days and 37% misclassified non-avalanche days compared to 33% and 31% in the tree analysis likely because the nearest neighbour model was optimized with a threshold of three nearest neighbour days. In the test sample 47% of avalanche days and 29% of non-avalanche days were misclassified in the nearest neighbour model using Sn38 compared to 31% misclassified avalanche days and 16% misclassified non-avalanche days in the tree analysis. It is expected that Cornice would perform better if the weights could be set using five nearest neighbours as the threshold for forecasting an avalanche day. The correct forecast of non-avalanche days and

avalanche days yielded almost the same accuracy in the learning sample with a fit of 63% when including Sn38 and 64% when excluding Sn38. More avalanche days were correctly forecast when Sn38 was used as predictor variable in the learning sample than when excluding Sn38 (64% compared to 60%) but not in the test sample where the forecast accuracy was 55% without Sn38 and 53% with Sn38. However non-avalanche days were forecast more accurately without Sn38 by 5%, 6% and 6% for (a), (b) and (c).

AvalDay(2) can be again be used in Cornice with the standard configuration (three nearest neighbours), because with 12% avalanche days the dataset is unbalanced. The results are summarized in Table 9.23.

Table 9.23. Performance of Nearest Neighbour model. AvalDay(2).

		Days forecast	*Unstable days (DSI = 1)*			*Stable days (DSI = 0)*			*Fitness of batch test (%)*
			Correct	*Wrong*	*% correct*	*Correct*	*Wrong*	*% correct*	
With Sn38	(a)	821	47	49	49	618	107	85	67
	(b)	237	10	22	31	184	21	90	61
	(c)	1065	60	69	47	783	153	84	65
Without Sn38	(a)	821	37	59	39	668	57	92	65
	(b)	237	13	19	41	187	18	91	66
	(c)	1065	52	77	40	857	79	92	66

The results are slightly better when Sn38 is included with an overall better forecast of 2%, 5% and 1% in (a), (b) and (c) respectively. However the results are not satisfactory for predicting avalanche days in any samples. By including Sn38, the forecast for avalanche days improved in the learning sample but not in the test sample. This was also observed for AvalDay(1) suggesting that the past two years must have been different from the other years in the dataset in regard to their meteorological predictor variables. The overall performance of all samples is between 61% and 67%, but the forecast of avalanche days was 31% to 49% correct with and without Sn38, respectively. This suggests that the definition of these categorical response variables may not be suitable for avalanche forecasting at Rogers Pass even though the forecast for stable days was good with 84% to 92%.

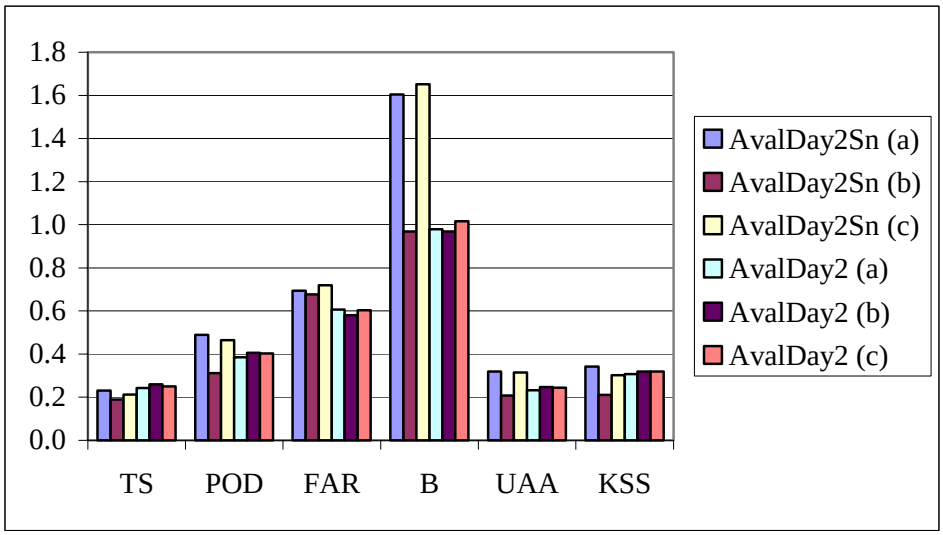

Figure 9.17: Forecast verification. AvalDay(2).

Figure 9.17 shows the corresponding accuracy measures and skill scores. The threat score (TS) are less than 0.3 which indicates that only a few avalanches days were correctly forecast even when considering the unbalanced dataset. This is also reflected in the poor score for the unweighted average accuracy (UAA). The false alarm rate (FAR) is around 0.6 when excluding Sn38 and even higher when including Sn38 indicating that avalanche days were forecast in more than 50% of the non-avalanche days. The Kuipers skill scores (KSS) ranges from 0.21 to 0.34 indicating the overall unsatisfactory forecast, however the results are slightly better when excluding Sn38. The bias (B) when including Sn38 shows that the learning sample was overforecast and the test sample slightly underforecast, which is in accordance with the poor result from the test sample for the other accuracy measures because less avalanche days were forecast than actually occurred.

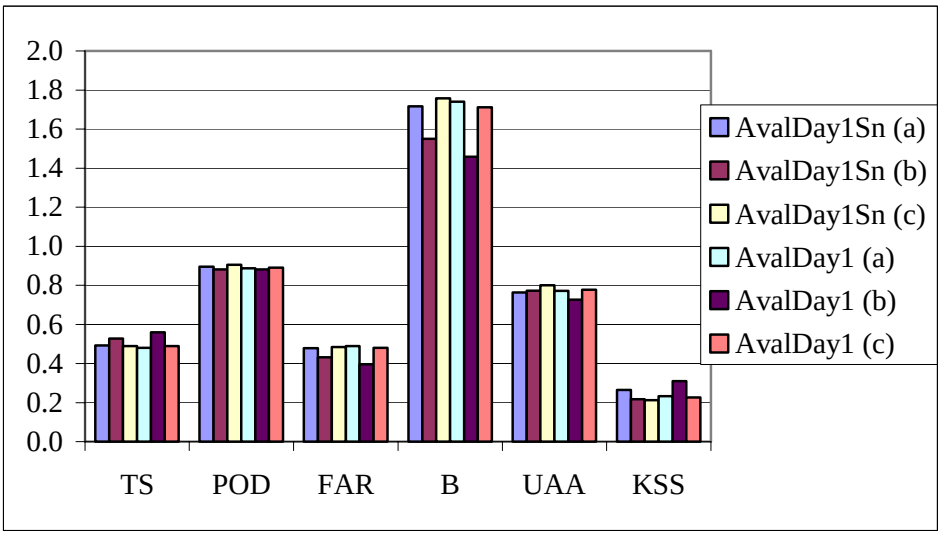

Figure 9.18: Forecast verification. AvalDay1 (3 NN).

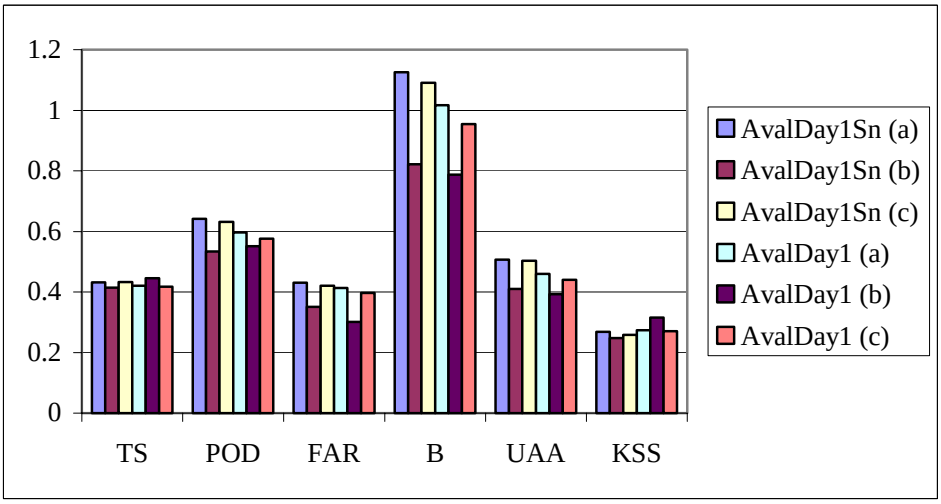

Figure 9.19: Forecast verification. AvalDay(1) (NN5).

Figure9.18 shows how carefully these results have to be interpreted and indicates the importance of chosing the threshold number of nearest neighbour days for forecasting an avalanche day. The accuracy measures on their own would indicate a good forecast with probability of detection (POD) almost 0.9 and an unweighted average accuracy (UAA) of almost 0.8. However the FAR is rather high but for avalanche forecasting purposes acceptable. The Kuipers skill score (KSS) indicates that the forecast was not much better

than random (less than 0.3). The bias indicates that avalanche days were overforecast in all samples.

The forecast verification in Figure 9.19 shows better results when using five nearest neighbours as the threshold. However false unstable predictions (FAR 0.3-0.4) are often made considering that 0 would be perfect but 0.3 to 0.4 is acceptable because the consequences are not as great as for false stable forecasts. The bias shows that in three out of six possibilities an underforecast occurred and in the other three cases a slight overforecast. The POD around 0.6 is reasonable showing that the fraction of correctly forecast avalanche days was about 60%. However the skill score KSS of less than 0.3 indicates that the forecast was not much better than a random forecast.

As already seen the tree analysis the suitability of the AvalDay(1) and AvalDay(2) is questionable and the problem more complex for the forecasting purpose.

9.8 Summary

In this Chapter the value of a stability index to forecast natural avalanche activity (Sn38) was assessed using several response variables. The shear frame measurements used for the calculations of Sn38 were provided by the avalanche control section in Glacier National Park who also provided the avalanche occurrence data. The response variables were defined in accordance to the forecast purpose of protecting the traffic on the highway. Size 1 and size 2 avalanches are unlikely to reach the highway and not considered a threat. Two definitions for an avalanche day were considered (AvalDay (1) and (2)) as well as the maximum size of an avalanche occurrence and an avalanche index. Spearman rank correlation analysis showed the strongest correlation with the precipitation over the past 24 hours in all datasets, which was also the most influential predictor variable in the classification tree analysis in all datasets.

Comparing the classification tree results of the two definitions of avalanche days shows that AvalDay (2) (a day with avalanches of size 3 or larger) was better predicted than AvalDay (1) which includes avalanches of size 2 or larger. These two datasets were the only ones included in a nearest neighbour model. However problems occurred due the balance of the datasets. AvalDays (1) consists of 40% avalanche days, which is a very high

percentage in contrast to other forecasting datasets. Cornice uses a threshold of three nearest neighbour days to forecast an avalanche day, but using this threshold in an almost balanced dataset leads to an overcast of avalanche days and in about 80% of the forecasts, an avalanche day was predicted. Considering a threshold of five nearest neighbours yielded a more realistic forecast of 62%-63% accuracy. The nearest neighbour forecast for AvalDay(2) was only more accurate including Sn38 in the learning but not in the test sample which was in accordance to the results from the classification tree analysis.

The classification tree analysis for the maximum sized avalanche and the avalanche index showed that Sn38 had predictive merit though not as dominantly as for the forecast of skier-triggered avalanches.

10. Conclusions

In this thesis an attempt was made to improve the forecast of skier-triggered avalanches by incorporating stability indices and snowpack properties as predictor variables into a daily avalanche forecasting model. This required that the shear strength of potential weak layers be extrapolated over time in order to be able to calculate daily stability indices. In addition the slab thickness and load were extrapolated on days without manual snowpack observations. Previous work at the University of Calgary showed that it is possible to extrapolate the shear strength of persistent weak layers over time and that a stability index correlated with the avalanche activity. However none of the studies on persistent weak layers adjusted the shear strength to the normal load. In Chapter 5 a normal load adjustment was found for the shear strength of persistent weak layers (Equation 5.3b) and in the analysis in this thesis the shear strength of a persistent weak layer was adjusted accordingly. The achievements in this thesis are, by Objective (Section 1.7):

Objective 1: two empirical models were developed to calculate the shear strength of layers of faceted crystals and surface hoar layers, respectively. The shear strength model for layers of faceted crystals is based on the load above a weak layer. Load has shown to be one of the most important predictor variables for shear strength, though the shear strength likely depends not only on load. However the limited dataset did not allow a multivariate regression analysis. The best results were achieved when applying a Power Law relationship, which fitted the data with an r^2 of 83% (Equation 6.3a). The surface hoar dataset was much larger and a more detailed analysis was possible. Logarithmic as well as multiple regression analysis showed promising results; however Equation 6.8, a multiple regression using the thickness of the weak layer, the temperature 5 cm below the weak layer, the snowpack height, the slab thickness and load, was chosen for further analysis. The r^2 was 0.83. Various models were assessed depending on the relevant weak layer depth for skier triggering, the coefficient of determination and the possible violations of assumptions in regression analysis.

Objective 2: the second objective built on the empirical models in Objective 1 but further the shear strength change rate was determined for estimating the shear strength on days without manual snowpack measurements. However the shear strength change rate could not be determined using regression analysis for layers of faceted crystals because of the limited dataset. It seemed to be more suitable to use long term average loading rates from the Columbia Mountains of Canada or daily loading rates measured with a precipitation gauge to determine the shear strength change rate because load was found to be the most influential predictor variable. Although there was a larger dataset for surface hoar, the coefficient of determination for any of the defined models yielded only 0.21 and consequently the multiple regression was deemed unsuitable for a predictive model. Again the average and the daily loading rates were the most promising.

Objective 3: In Equation 6.1 an Interval Model was introduced which combines the results of Objective 1 and 2 into a Forecasting Model to calculate the shear strength of persistent weak layers between days with manual snowpack observations. In Objective 2 it was suggested to use average and daily loading in the calculations for the shear strength change rate. During the analysis it became clear that daily loading rates seemed to be a better predictor because above and below average snow years are better forecast and also responds better to specific weather patterns. Consequently Equation 6.6a and Equation 6.1a were used to extrapolate the shear strength over time, which was then used to calculate the skier stability index Sk38. It is also possible to estimate the shear strength from rutschblock test results and Sk38 was calculated using these results. Further analysis included both shear strength estimations.

Objective 4, 7, 8: In Chapter 7 several datasets were analysed to assess the importance of daily stability indices and snowpack properties in forecasting skier-triggered avalanches on persistent weak layers. The analysis included Spearman rank correlations, classification tree analysis and nearest neighbour analysis. Sk38, Load and *H* clearly improved the forecast of skier-triggered avalanches in all the datasets under consideration. The incorporation of results from rutschblock tests was

unsatisfactory and it was not justifiable to use the results in a regional forecasting model.

Objective 5, 7, 8: Even though persistent weak layers are of greater concern to skiers, skier triggering on non-persistent weak layers was assessed to be able to run a daily forecasting model in the future. Shear frame measurements were available in one study area but only once per day whereas the shear strength of storm snow layers often changes within hours. It was not possible to formulate an empirical model to forecast the shear strength of non-persistent weak layers and consequently the analysis was based on measured values. An attempt to forecast storm snow avalanches in times during and soon after a storm resulted in poor predictions, but a daily model to predict non-persistent weak layers on a daily basis was more suitable for the purpose of this thesis. However using the shear frame measurements did not improve the forecast. Nevertheless, the results showed some potential to predict avalanches and it is suggested that a refined shear strength model for storm snow may improve the forecast in the future.

Objective 6: Shear frame tests have been used for stability evaluations of natural avalanche activity in storm snow for a long time and even though skier-triggered avalanches were the focus of this study it seemed to be worthwhile to calculate a regional stability index for natural avalanche (Sn38) to assess its predictive value. The shear frame measurements were provided by a forecasting service for a highway and it was attempted to assess the value of these measurements in accordance to their forecasting purpose. Consequently several response variables were assessed mostly based on the critical size of an avalanche reaching the highway. The value of the shear frame tests was not as promising as expected likely because none of the response variables were well suited to forecasting for avalanches that threaten the highway.

All in all the analysis showed that stability indices and snowpack properties have predictive value to forecast skier-triggered avalanches. The best results were achieved for persistent weak layers, which was desirable because they pose the greatest threat to skiers. However the forecasts for non-persistent weak layers lacked accuracy

and needs further analysis. The results for persistent weak layers are promising and may help the forecaster in the decision-making process as a tool to analyse and visualize complex datasets. The nearest neighbour approach seems to be a good choice, because the statistical analysis is straightforward, easy to apply and models are already commercially available. Also, the list of avalanches that occurred on similar days is of great value to forecasters. The results from a classification tree yielded similar results to the nearest neighbour model, but the practical applicability in forecasting operations is up to now limited. The results for natural avalanche activity could not be improved by using a stability index and consequently a forecasting model based on a yes/no forecast seems to be not suitable in the study area analysed. Even though the results of a nearest neighbour model to forecast skier-triggered avalanches are good, the forecasters should analyse the results carefully and understand what a computer assisted forecasting model can offer and what it cannot offer.

Apart from the achievements related to the objectives, the results contribute to avalanche forecasting by providing a tool for daily stability assessments. In most backcountry skiing operations the stability is assessed using conventional avalanche forecasting methods as described in LaChapelle (1980) or by combining statistical and causal-intuitive approaches which is, as LaChapelle (1970) said, likely the method of the future. In most avalanche studies involving statistical methods, a historical database has been used to analyse the stability in a forecast area. The nearest neighbour approach was chosen in this study, because this method is easy to understand and is similar to conventional avalanche forecasting procedures. The principle of the nearest neighbour avalanche forecast fits into the fundamental processes in conventional avalanche forecasting as described by LaChapelle (1980) as a mix of deterministic treatment for snow and weather parameters and inductive logic to reach actual forecast decisions. In general, some of the same predictor variables are used to assess the stability on a forecast day and similar situations are displayed, which can complement the memory of the forecaster. The information a regional model can give does not account for specifics in the area but gives a hypothesis about the snow stability using the initially available data (LaChapelle, 1980) and can then be revised according to the observations in the field. However in conventional

forecasting it is not common to make calculations such as are required for skier stability indices; rather, information gathered in the field is used to determine the condition of weak layers, though a computer model can help to analyse this information.

Another advantage is that information of the previous days is easily assessable in a nearest neighbour model and can be used by returning staff without breaking the iteration process which was observed by LaChapelle (1980) as a source of error because of difficulties in re-establishing the base of prior knowledge, which is essential for the evaluation of the snow stability.

Other computer forecasting models have been developed, including Davis and Elder (1995) classification trees, McClung and Tweedy (1993, 1994) discriminant analysis at Kooteney Pass, Floyer's (2003) results at Bear Pass, Buser's (1989) nearest neighbour model in Switzerland and the ones introduced in this study. While these other studies agree with the current study with regard to the importance of variables for recent weather and avalanche activity, this is the first study to assess the importance of a skier stability index as well as slab properties such as load and slab thickness over a particular weak layer. The importance of monitoring persistent weak layers for forecasting human triggered avalanches in the Columbia Mountains is consistent with Hägeli and McClung (2003) and Hägeli (2004). A direct comparison of forecasting accuracy from this study to the accuracy from the studies by Davis and others (1996), McClung and Tweedy (1994) and Floyer and McClung (2003) cannot be made because this thesis focuses on skier triggering rather than natural avalanching or explosive triggering. However it was shown that a stability index for natural avalanche activity was not as successful as for skier triggering for the same forecast area. The variables most significant in all these studies depended on the variable selection, the avalanches forecasted (dry, wet, natural, skier-triggered, etc.) and therefore each variable cannot be analysed. However it seems that in some areas the wind effect was greater (McCollister, 2004) whereas in others precipitation variables such as the new snow height (Floyer, 2003: 61) were more significant.

In the winter of 2003/2004 the model was tested in a trial at a Helicopter Skiing operation in Blue River. The operation expressed interest in the use of such a Nearest Neighbour model and in the coming winter the model will be run daily as a trial. However,

the snowpack variables Sk38, Load and H cannot be calculated in Cornice and have to be prepared before data entry. The model is not combined with a GIS and therefore the event list of the ten nearest neighbour days is basic and does not use visualization of the distribution of the avalanches on similar days. Additionally skier-triggered avalanches on persistent and non-persistent weak layers are best forecast with different predictor variables (Section 7.3 and Section 8.8) and Cornice has to be configured separately for these two forecasts. Nevertheless, this trail represents an important step in systematically using historical data to forecast skier-triggered avalanches.

11. Recommendations for further research

It has been shown that stability indices and snowpack properties have the potential to improve the computer assisted forecasts of avalanche activity on persistent weak layers. Nevertheless, the following research might improve avalanche forecasting, including computer assisted forecasting:

- continuous testing of near surface facets and facets on and below crusts in a controlled environment or the field to increase the dataset and find an empirical or physically based strength change model not based on load alone
- further assessment of the value of rutschblock tests and compression tests in regional avalanche forecasting by analysing whether the study site is representative or not
- estimation of the stability index Sk using rutschblock test results based on the normal load adjustment for persistent weak layers
- incorporation of fracture character (shear quality) and release type because it has been shown that the fracture character improves the interpretation of the stability test results
- shear frame measurements on storm snow should be recorded more consistently and the failure layer monitored over time
- shear frame measurements on storm snow several times per days in a controlled environment or during storm periods in order to be able to model shear strength over time
- better avalanche occurrence data clearly indicating the failure layer for each avalanche and whether no avalanches were recorded because none occurred or observations were impossible
- Refinement of Cornice based on the optimum number of nearest neighbours as threshold to determine an avalanche day

- Systematic analysis of misclassified cases to find similar patterns, for instance thick crusts which than can be considered to be added as additional predictor variable
- Refinement of DSI based on more consistent avalanche occurrence data
- More analysis of the potential predictors for natural avalanches using hourly rather than daily values for the predictors and avalanche activity.

References

Allen, M.P. 1997. *Understanding regression analysis*. Plenum Press. New York. 216 pp.

Armstrong, R.L. and Armstrong, B.R., 1987. Snow and avalanche climates of the western United States: A comparison of maritime, intermountain and continental conditions. Avalanche Formation, Movement and Effects. *Proceedings of the Davos Symposium*, September 1986. IAHS Publ. No. 162. 281-294.

Atkins, R., 1992. Computer graphics applications in avalanche forecasting. *Proceedings of the International Snow Science workshop*. 116-125.

Bader, H., Haefeli R., Bucher, E., Neher, J. Eckel, O. and Thams, C., 1939. *Der Schnee und seine Metamorphose. Beiträge zur Geologie der Schweiz*. Schweizerische Schnee- und Lawinenforschungskommission. Translation: 1954. Snow and its metamorphism. SIPRE Translation 14.

Ballard, G.E.H. and Feldt, E.D., 1965. A theoretical consideration of the strength of snow. *Journal of Glaciology*, Vol. 6(43)**.** 159-170.

Birkeland, K.W., 1998. Terminology and predominant processes associated with the formation of weak layers of near-surface faceted crystals in the mountain snowpack. *Arctic and Alpine Research*, Vol. **30** (2). 193-199.

Birkeland, K. and Johnson, R., 1996. The stuffblock snow stability test. *6E62A95*, USDA Forest Service, Technology and Development Center Missoula, Missoula MT, U.S.A.

Birkeland, K.W., Johnson, R.F. and Schmidt, D.S., 1998. Near-surface faceted crystals formed by diurnal recrystallization: a case study of weak layer formation in the mountain snowpack and its contribution to snow avalanches. *Arctic and Alpine Research*, Vol. **30** (2). 200-204.

Bois, P., Obled, C. and Good, W., 1974. Multivariate data analysis as a tool for day-by-day avalanche forecast. Snow mechanics symposium. *IAHS-AISH Publication No. 114. Proceedings of the Grindelwald symposium, April 1974*. 391-403.

Bolognesi, R., 1992. Artificial intelligence and local avalanche forecasting: the system „Avalog“. *International Emergency and Engineering conference, Arlington.* 113-116.

Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J., 1984. *Classification and Regression Trees*. Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA.

Brun, E., David, P., Sudul, M. and Brunot, G., 1992. A numerical model to simulate snow-cover stratigraphy for operational avalanche forecasting. *Journal of Glaciology*, Vol. **38** (128). 13-22.

Brun, E., Martin, E., Simon, V., Gendre, C. and Coleou, C., 1989. An energy and mass model of snow cover suitable for operational avalanche forecasting. *Journal of Glaciology*, Vol. **35** (121). 333-342.

Bucher, E., 1948. Beitrag zu den theoretischen Grundlagen des Lawinenverbaus. *Beitrage zur Geologie der Schweiz. Geotechnische Serie. Hydrologie*, Lief. 6.

Burt, J.E. and Barber, G.M., 1996. *Elementary statistics for geographers.* Second edition. New York. 640 pp.

Buser, O., 1989. Two years experience of operational avalanche forecasting using the nearest neighbours method. *Annals of Glaciology*, Vol. **13**. 31-34.

Buser, O., Föhn, P., Good, W., Gubler, H. and Salm, B., 1985. Different methods for the assessment of avalanche danger. *Cold Regions Science and Technology*, Vol. **10**. 199-218.

Canadian Avalanche Association (CAA), 1995. *Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches*. Canadian Avalanche Association, Revelstoke, BC.

Canadian Avalanche Association (CAA), 2002. *Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches*. Canadian Avalanche Association, Revelstoke, BC.

Campbell, C.P., 2004. *Spatial variability of slab stability and fracture properties in avalanche start zones.* MSc thesis. University of Calgary.

Chalmers, T., 2001. *Forecasting shear strength and skier-triggered avalanches for buried surface hoar layers.* MSc thesis. University of Calgary.

Chalmers, T. and Jamieson, B., 2002. Forecasting shear strength and skier-triggered avalanches for buried surface hoar layers. *Presented at the 2002 International Snow Science Workshop in Penticton, Canada.* Snow Avalanche Programs, BC Ministry of Transportation, Victoria, Canada.

Chalmers, T. and Jamieson, B., 2003. A snow-profile-based forecasting model for skier-triggered avalanches on surface hoar layers in the Columbia Mountains of Canada. *Cold Regions Science and Technology*, Special edition.

Colbeck, S.C., 1991. The layered character of snow covers. *Reviews of Geophysics*, Vol. **29** (1). 81-96.

Colbeck, S.C., 1997. A review of sintering in seasonal snow. *CRREL Report* 97-10. 11pp.

Colbeck, S.C. 1998. Sintering in a dry snow cover. Journal of Applied Physics 84(8), 4585-4589.

Colbeck, S.C. and Jamieson, J.B., 2001. The formation of faceted layers above crusts. *Cold Regions Science and Technology*, Vol. **33**. 247-252.

Colbeck, S., Akitaya, E., Armstrong, R., Gubler, H., Lafeuille, J., Lied, K., McClung, D. and Morris, E., 1990. *The International Classification for Seasonal Snow on the Ground.* Issued by The International Commission on Snow and Ice of the International Association of Scientific Hydrology and Co-Issued by: Inter. Glaciological society.

Conway, H. and Abrahamson, J., 1984. Snow stability index. *Journal of Glaciology*, Vol. **30** (116). 321-327.

Conway, H. and Wilbour, C., 1999. Evolution of snow slope stability during storms. *Cold Regions Science and Technology*, Vol. **30**. 67-77.

Davis, R.E. and K. Elder., 1995: Application of classification and regression trees: selection of avalanche activity indices at Mammoth Mountain. *Proceedings 1994 International Snow Science Workshop. ISSW'94,* P.O. Box 49, Snowbird, Utah, USA. 285-294.

Davis, R.E., Elder, K., Howlett, D. and Bouzaglou, E., 1996. Analysis of weather and avalanche records from Alta, Utah and Mammoth Mountain, California using classification trees. *Proceedings 1996 International Snow Science Workshop. ISSW'96,* Canadian Avalanche Centre, P.O. Box 2759, Revelstoke, BC V0E 2S0, Canada. 14-19.

deQuervain, M., 1950. Die Festigkeitseigenschaften der Schneedecke und ihre Messung (strength properties of snowpacks and measurements). *Geofisica Pura E Applicata*, Vol. **18**. 3-15.

Duclos, A., 1998. Slab avalanches and “new” snow. *Proceedings of the International Snow Science Workshop. ISSW'98,* Canadian Avalanche Centre, P.O. Box 2759, Revelstoke, BC V0E 2S0, Canada. 478-488.

Durand, Y., Giraud, G., Brun, E., Mérindol, L. and Martin, E., 1999. A computer-based system simulating snowpack structures as a tool for regional avalanche forecasting. Journal of Glaciology, Vol. 45 (151). 469-484.

Edens, M.Q. and Brown, R.L., 1991. Changes in microstructure of snow under large deformations. Journal of Glaciology, Vol. 37 (126). 193-202.

Farrer, D.E. and Glauber, R.R., 1967. Multicollinearity in regression analysis: the problem revisited. Review of Economics and Statistics, 49, 92-107.

Feldt, E.D. and Ballard G.E.H., 1966. A theory of the consolidation of snow. Journal of Glaciology, Vol. 6 (43). 145–157.

Ferguson, S.A., Moore, M.B., Marriott, R.T. and Speers-Hayes P., 1990. Avalanche weather forecasting at the northwest avalanche center, Seattle, Washington, USA. Journal of Glaciology, Vol. 36 (122). 57-66.

Fierz, C.H., 1998. Field observation and modelling of weak-layer evolution. Annals of Glaciology, Vol. 26. 7-13.

Floyer, J.A., 2003. Statistical avalanche forecasting using meteorological data from Bear Pass, British Columbia, Canada. MSc Thesis. University of British Columbia. 98pp.

Föhn, P.M.B., 1987a. The stability index and various triggering mechanisms. In: Salm, B. and Gubler, H., (eds.), Avalanche Formation, Movement and Effects. International Association of Hydrological Sciences, Publication No. 162. 195-211.

Föhn, P.M.B., 1987b. The rutschblock as a practical tool for slope stability evaluation. In: B. Salm and H. Gubler, eds., Avalanche Formation, Movement and Effects. International Association of Hydrological Sciences, Publication No. 162. 223-228.

Föhn, P.M.B. and Camponovo, C., 1997. Improvements by measuring the shear strength of weak layers. In ISSW '96. International Snow Science Workshop, 6-10 October 1996, Banff, Alberta. Proceedings. Revelstoke, B.C., Canadian Avalanche Association. 158-162.

Fukuzawa, T. and Akitaya, E., 1993. Depth-hoar crystal growth in the surface layer under high temperature gradient. Annals of Glaciology, Vol. 18. 39-45.

Garson, G.D., n.d (retrieved November 2004). Statsnotes: An online textbook. Multiple regression. http://www2.chass.ncsu.edu/garson/pa765/statnote.htm.

Gassner, M., Birkeland, K., Etter, H.J. and Leonard, T., 2000. An improved avalanche forecasting program based on the nearest neighbor method. In *International Snow Science Workshop*, Big Sky, Montana, U.S.A., 2000.

Giraud, G., 1992. Mepra an expert system for avalanche risk forecasting. Proceeding of the International Snow Science Workshop. 97-104.

Gubler, H. 1982. Strength of bonds between ice grains after short contact times. Journal of Glaciology, Vol. 28 (100). 457-473.

Haefeli, R., 1967. Some mechanical aspects on the formation of avalanches. Physics of Snow and Ice, (H. Oura, Ed.) Vol. 1, Part 2, Institute of Low Temperature Science, Hokkaido University, Japan. 119-1213.

Hägeli, P. and McClung, D.M., 2001. A new perspective on computer-assisted avalanche forecasting: Scale and scale issues. In ISSW '00. International Snow Science Workshop, 1-6 October 2000, Big Sky, Montana. Proceedings. American Avalanche Association. 66-73.

Hägeli, P. and McClung, D. M., 2002. Analysis of weak layer avalanche activity in the Columbia Mountains, British Columbia, Canada. In Stevens, J.R. (ed.), International Snow Science Workshop 2002, 29 September—4 October 2002, Penticton, British Columbia, Proceedings. Victoria, B.C., B.C. Ministry of Transportation, Snow Avalanche Programs, 1-7.

Hägeli, P., McClung, D.M., 2003. Avalanche characteristics of a transitional snow climate – Columbia Mountains, British Columbia, Canada. Cold Reg. Sci. Technol. 37(3), 255-276.

Heierli. J., Purves, R.S., Felber, A. and Kowalski, J., 2004. Verification of nearest neighbours interpretations in avalanche forecasting. Annals of Glaciology, Vol. 38. In press.

Hillel, D., 1980. Fundamentals of soil physics. Academic Press, New York.

Jamieson, J.B., 1995. Avalanche prediction for persistent snow slabs. PhD Thesis, University of Calgary.

Jamieson, J.B., 1999. The compression test - after 25 years. The Avalanche Review, Vol. 18(1). 10-12.

Jamieson, J.B. and Geldsetzer, T., 1996. Avalanche accidents in Canada. Volume 4 1984-1996. Canadian Avalanche Association. Revelstoke. 193pp.

Jamieson, J.B. and Johnston, C.D., 1990. In situ tensile tests of snowpack layers. Journal of Glaciology, Vol. 36(122). 102-106.

Jamieson, J.B. and Johnston, C.D., 1993. Shear frame stability parameters for large scale avalanche forecasting. Annals of Glaciology, Vol. 18. 268-273.

Jamieson, J.B. and Johnston, C.D., 1997. The facet layer of November 1996. Avalanche News 52, 10-15.

Jamieson, J.B. and Johnston, C.D., 1998. Refinements to the stability index for skier-triggered dry-slab avalanches. Annals of Glaciology, Vol. 26. 296-302.

Jamieson, J.B. and Johnston, C.D., 2001. Evaluation of the shear frame test for weak snowpack layers. Annals of Glaciology, Vol. 32. 59-69.

Jamieson, J.B. and van Herwijnen, A., 2002. Preliminary results from controlled experiments on the growth of faceted crystals above a wet snow layer. Proceedings International Snow Science Workshop, Penticton BC, Canada, 29 September - 4 October 2002.

Jamieson, B., Geldsetzer, T. and Stethem, C., 2001. Forecasting for deep slab avalanches. Cold Regions Science and Technology, Vol. 33(2-3). 275-290.

Jarrett, J. and Kraft, A., 1989. Statistical analysis for decision making. Allyn and Bacon, Needham Heights, MA.

Johnson, G.T., 2000. *Observations of Faceted Crystals in Alpine Snowpacks*. MSc Thesis. Dept. of Civil Engineering, University of Calgary, Calgary, Canada. 98 pp.

Johnson, R.A. and Bhattacharyya, G.K., 1996. *Statistics. Principles and Methods*. 3rd edition, John Wiley & Sons, Inc., US.

Jones, A.S.T. and Jamieson, J.B., 2001. Meteorological forecasting variables associated with skier-triggered dry slab avalanches. *Cold Regions Science and Technology*, Vol. **33**. 223-236.

Judd, C.M. and Clelland, G.H., 1989. *Data analysis. A model-comparison approach*. Orlando.

Judson, A. and Erickson, B.J., 1973. *Predicting avalanche intensity from weather data: a statistical analysis.* USDA Forest Service. Research Paper RM-112. 12p.

Keeler, C. M., 1969. Some physical properties of alpine snow. CRREL Research Report 271. Cold regions research and engineering laboratory. 69 pp.

Keeler, C.M. and Weeks, W.F., 1967. Some mechanical properties of alpine snow, Montana 1964-66. CRREL Research Report 227. Cold regions research and engineering laboratory. 43 pp.

Kojima, K., 1954. Viscoelastic property of snow. Low Temperature Science, Series A. Vol. 12. 1–13.

Kojima, K., 1967. Densification of seasonal snow cover. In: Oura H. (ed.) Physics of Snow and Ice Vol. I Part 2, Proceedings of a Conference on Physics of Snow and Ice, Institute of Low Temperature Science, Hokkaido University. 929-952.

LaChapelle, E. R., 1966. Avalanche Forecasting--a Modern Synthesis. *Proceedings of Davos Symposium.* IASH Publ. No. 69. p. 350-356.

LaChapelle, E. R., 1970. Principles of Avalanche Forecasting. In: Gold, L. W. and Williams, G. P. (eds.), *Ice Engineering and Avalanche Forecasting and Control*. NRC of Canada, Assoc. Comm. On Geotechnical Research, Tech. Memoir. No. 98. 106-113.

LaChapelle, E.R., 1980. The fundamental processes in conventional avalanche forecasting. *Journal of Glaciology*, Vol. **26** (94). 75-84.

Landry, C., Birkeland, K., Hansen, K., Borkowski, J., Brown, R. and Aspinall, R., 2002. Snow stability on uniform slopes: implications for avalanche forecasting. In *ISSW '02. International Snow Science Workshop, 29 September – 4 October, Penticton. Proceedings.* BC Ministry of Transportation, Victoria, BC. 532-539.

Laternser, M. and Schneebeli, M., 2002. Temporal trend and spatial distribution of avalanche activity during the last 50 years in Switzerland. *Natural Hazard*, **27**. 201-230.

Lehning, M., Bartelt, P., Brown, B., Russi, T., Stöckli, U. and Zimmerli, M., 1999. Snowpack model calculations for avalanche warning based upon a new network of weather and snow stations. *Cold Regions Science and Technology*, Vol. **30**. 145-157.

Manley, B.F.J, 1994. *Multivariate statistical methods. A primer.* Second edition. Chapman & Hall, London.

Marshall, T.J. and Holmes, J.W., 1979. *Soil physics*. Cambridge University Press, Cambridge.

McClung, D.M., 1995. Use of expert knowledge in avalanche forecasting. *Defence Science Journal*, Vol. **45** (2). 117-123.

McClung, D. and Schaerer, P.A., 1993. *The Avalanche Handbook*. The Mountaineers, Seattle.

McClung, D. and Tweedy, J., 1994. Numerical avalanche prediction: Kootenay Pass, British Columbia, Canada. *Journal of Glaciology*, Vol. **40** (135). 350-358.

McCollister, C.M., 2004. *Geographic knowledge discovery techniques for exploring historical weather and avalanche data.* MSc Thesis. Montana State University.

Mendenhall, W. and Sincich, T., 1996. *A second course in statistics. Regression analysis*. 5th edition. New Jersey.

Mérindol, L., Guyomarc'h, G. and Giraud, G., 2002. A French local tool for avalanche hazard forecasting: Astral, current state and new developments. *Proceedings of the International Snow Science Workshop, Pentiction, Canada*. 105-108.

Mock, C.J. and Kay, P.A., 1992. Avalanche climatology of the western United States, with an emphasis on Alta, Utah. *The Professional Geographer*, Vol. **44**. 307-318.

Perla, R.I., 1970. On contributory factors in avalanche hazard evaluation. *Canadian Geotechnical Journal*, Vol. **7**. 414-419.

Perla, R.I. 1977. Slab avalanche measurements. Canadian Geotechnical Journal 14(2), 206-213.

Perla, R.I., 1970. On Contributory Factors in Avalanche Hazard Evaluation. *Canadian Geotechnical Journal*, Vol. **7** (4). 414-419.

Purves, R.S., Morrison, K.W., Moss, G. and Wright, D.S.B., 2003. Nearest neighbours for avalanche forecasting in Scotland – development, verification and optimisation of a model. *Cold Regions Science and Technology*, Vol. **37**, special edition. 343-355.

Roch, A., 1965. Les variations de la resistance de la neige. *International symposium on scientific aspects of snow and ice avalanches*. AIHS-Publication 69. 86-99.

Roch, A., 1966. Les declenchements d'avalanches. *Union de Geodesie et Geophysique Internationale. Association Internationale d'Hydrologie Scientifique.* Commission pour la Neige et la Glace. Division Neige Saisonniere et Avalanches. *Symposium international sur les aspects scientifiques des avalanches de neige,* 5-10 Avril 1965, Davos, Suisse. 182-195.

Rosenthal, W. and Elder, K., 2002. Evidence of chaos in slab avalanching. *Proceedings of the 2002 International Snow Science Workshop*, Penticton, British Columbia, Canada, September 2002.

Russi, T., Ammann, W., Brabec, B., Lehning, M. and Meister, R., 2000. Avalanche Warning Switzerland 2000. In *International Snow Science Workshop 1998, Proceedings,* Sunriver, Oregon, U.S.A. 146-153.

Salway, A.A., 1979. Time-series modelling of avalanche activity from meteorological data. *Journal of Glaciology*, Vol. **22** (88). 513-528.

Schleiss, V. G. and Schleiss, W. E., 1970. Avalanche Hazard Evaluation and Forecast, Rogers Pass, Glacier National Park. *NRC of Canada Tech. Memoir*. No. 98. 115-122.

Schleiss, V.G., 1989. *Rogers Pass Snow Avalanche Atlas, Glacier National Park, British Columbia, Canada*. Parks Canada, Revelstoke, BC.

Schweizer, J. and Camponovo, C., 2001. The skier's zone of influence in triggering slab avalanches. *Annals of Glaciology*, Vol. **32.** 314-320.

Schweizer, J. and Föhn, P.M.B., 1996. Avalanche forecasting – an expert system approach. *Journal of Glaciology*, Vol. **42** (141). 318-332.

Schweizer, J. and B. Jamieson. 2001. Snow cover properties for skier-triggered avalanches. *Cold Regions Science and Technology* **33** (2-3). 207-221.

Schweizer, J., Kronholm, K., 2004. Snow cover spatial variability at multiple scales. *Proceeding of the International Snow Science Workshop*, Jackson, Wyoming, USA, Sept. 19-24, 2004.

Schweizer, M., Föhn, P.M.B. and Schweizer, J., 1994. Integrating neural networks and rule based systems to build an avalanche forecasting system. *Proceedings of the IASTED International Conference Artificial intelligence, expert systems and neural networks*. Zürich, Switzerland July 4-6, 1994.

Schweizer, J., Jamieson, J.B. and Schneebeli, M., 2003. Snow avalanche formation. *Reviews of Geophysics*, Vol. **41** (4). 2-1 –2-11.

Sommerfeld, R.A., 1973. Statistical problems in snow mechanics. *U.S. Forest Service General Technical Report* RM-3. 29-36.

Sommerfeld, RA., King, R.M. and Budding, F., 1976. A correction factor for Roch's stability index of slab avalanche release. *Journal of Glaciology*, Vol. **17**.145–147.

Sommerfeld, R.A., 1984. Instructions for using the 250 cm^2 shear frame to evaluate the strength of a buried snow surface. *USDA Forest Service Research Note* RM-446. 1-6.

Sommerfeld, R.A. and King, R.M., 1979. A recommendation for the application of the Roch index for slab avalanche release. *Journal of Glaciology*, Vol. **22** (88). 547-549.

Stephens, J. E. Adams, X. Huo, J. Dent, J. Hicks and D. McCarty. 1995. Use of neural networks in avalanche hazard forecasting. Proceedings of the 1994 International Snow Science Workshop in Snowbird, Utah. ISSW '94, P.O. Box 49, Snowbird, Utah, USA, 327-340.

Stethem, Ch., Jamieson, B., Schaerer, P., Liverman, D., Germain, D. and Walker, S., 2003. Snow avalanche hazard in Canada – a review. *Natural Hazards*, Vol. **28.** 487-515.

Stewart, K., 2002. *Spatial variability of snow stability within avalanche start zones*. MSc Thesis. University of Calgary.

Stoffel, A., Meister, R. And Schweizer, J., 1998. spatial characteristics of avalanche activity in an alpine valley – a GIS approach. *Annals of Glaciology*, Vol. **26**. 329-336.

Wilks, D. S., 1995. *Statistical Methods in the Atmospheric Sciences*, Academic Press, 467pp.

Yoshida, Z., 1955. The Institute of Low Temperature Science, Hokkaido University, Japan. *Journal of Glaciology*, Vol. **2** (17).

Yoshida, Z., 1963. Physical properties of snow. In: Kingery, W.D. (ed.), *Ice and Snow: Properties, Processes and Applications*, Cambridge, Massachusetts: M.I.T Press. 485–527.

Zeidler, A. and Jamieson, J.B., 2002. Estimating the strength of faceted snow layers for an avalanche forecasting model. In: Stevens, J.R. (ed.), *International Snow Science Workshop 2002, 29 September—4 October 2002, Penticton, British Columbia, Proceedings.* Victoria, B.C., B.C. Ministry of Transportation, Snow Avalanche Programs. 130-137.

Zeidler, A. and Jamieson, J.B., 2004. A nearest-neighbour model for forecasting skier-triggered dry-slab avalanches on persistent weak layers in the Columbia Mountains, Canada. *Annals of Glaciology*, Vol. **38**. *In press.*

Table A1: Measured variables for model testing: Mt. St. Anne, faceted layer formed 07 January 2002 (Section 6.3.9).

Date	Age (days)	E min (mm)	Thick (cm)	T-5 (°C)	Ta (°C)	HS (cm)	H (cm)	Slab Dens (kg/m^3)	HH	TG (°C/ m)	Ta/HS (°C/m)	Load (kPa)	Strength (kPa)	Pcp (kPa)	Change (kPa/ day)
12-Jan-02	5	1.0	0.3	-4.1	-3.9	210	19	122	0.63	0.4	-0.02	0.23	0.33	-	0.103
15-Jan-02	8	0.3	0.3	-7.5	-11.0	236	19	140	1.00	-0.9	-0.05	0.26	0.63	0.01	0.001
18-Jan-02	11	0.5	0.3	-8.9	-11.5	222	21	131	1.00	-1.0	-0.05	0.27	0.64	0.01	0.057
22-Jan-02	15	0.5	0.5	-6.3	-10.7	250	49	112	2.53	-0.6	-0.04	0.53	0.87	0.27	0.068
29-Jan-02	22	1.0	0.3	-6.6	-15.4	281	57	189	4.00	-1.1	-0.05	1.14	1.34	0.24	0.082
02-Feb-02	26	1.0	0.3	-6.9	-5.1	260	76	183	6.32	-0.6	-0.02	1.40	1.67	0.10	0.012
05-Feb-02	29	2.0	0.3	-6.1	-4.5	280	74	189	2.53	-0.4	-0.02	1.31	1.71	0.08	0.061
19-Feb-02	43	2.0	0.4	-4.0	-3.9	293	110	215	10.13	-0.2	-0.01	2.15	2.57	0.24	0.013
26-Feb-02	50	1.5	0.4	-3.8	-8.5	310	119	230	6.32	-0.2	-0.03	2.59	2.66	0.28	0.090
05-Mar-02	57	1.0	0.5	-3.9	-18.9	296	115	250	40.50	-0.1	-0.06	2.77	3.29	0.17	0.191
13-Mar-02	65	1.0	1.5	-3.9	-1.0	336	153	228	16.00	0.2	0.00	3.34	4.82	0.33	-0.028
26-Mar-02	78	1.0	1.0	-3.7	-3.6	305	130	293	16.00	-0.1	-0.01	3.78	4.45	0.14	-

Table A2: Measured variables for model testing: Mt. St. Anne, surface hoar formed 10 February 1997 (Section 6.4.9).

Date	Age (days)	Emin (mm)	Thick (cm)	T-5 (°C)	HS (cm)	H (cm)	Load (kPa)	TG (°C/m)	SlabDens (kg/m^3)	Strength (kPa)	Change (kPa/day)	Pcp (kPa)
14-Feb-97	4	3.0	0.5	-5.9	289	30	0.26	-1.6	87.30	0.27	0.11	-
20-Feb-97	10	3.0	1.0	-5.2	315	67	1.23	0.3	187.97	0.92	0.26	0.40
24-Feb-97	14	6.0	0.5	-5.3	283	60	1.12	-0.1	190.48	1.96	0.16	0.11
01-Mar-97	19	1.0	0.5	-5.8	305	61	1.64	-0.8	273.22	2.78	0.08	0.12
05-Mar-97	23	1.0	0.3	-5.4	275	86	1.52	-0.4	179.96	3.09	-0.14	0.43
08-Mar-97	26	1.0	0.3	-5.2	355	116	1.96	-0.1	172.41	2.68	0.27	0.37
15-Mar-97	33	2.0	0.3	-4.6	338	106	2.24	-0.3	215.63	4.57	-0.01	0.29
22-Mar-97	40	1.0	0.3	-3.2	338	125	2.80	-0.1	228.57	4.50	0.08	0.49
27-Mar-97	45	3.0	0.4	-3.2	335	121	3.48	0.0	293.19	4.89	-	0.23

Table A3: Measured variables for model testing: Mt. Fidelity, surface hoar formed 28 January 2001 (Section 6.4.9).

Date	Age (days)	Emin (mm)	Thick (cm)	T-5 (°C)	HS (cm)	H (cm)	Load (kPa)	TG (°C/m)	SlabDens (kg/m3)	Strength (kPa)	Change (kPa/day)	Pcp (kPa)
30-Jan-01	2	5.0	1.2	-6.5	179	34	0.27	-0.7	80.28	0.39	0.08	-
02-Feb-01	5	4.0	0.6	-4.2	172	48	0.42	0.0	91.17	0.62	0.06	0.21
06-Feb-01	9	5.0	0.8	-4.2	212	92	0.83	0.1	91.77	0.86	0.22	0.44
10-Feb-01	13	3.0	0.4	-4.2	180	67	0.82	-0.6	124.60	1.73	-0.05	0.01
14-Feb-01	17	6.0	0.8	-3.8	163	54	0.84	-0.3	158.73	1.53	0.10	0.03
18-Feb-01	21	5.0	3.0	-5.7	179	62	1.03	-0.8	169.38	1.92	-0.06	0.08
22-Feb-01	25	5.0	0.6	-3.9	166	60	1.03	-0.2	174.85	1.68	0.08	0.09
26-Feb-01	29	4.0	0.5	-3.5	210	83	1.44	-0.5	176.50	2.02	0.18	0.18
01-Mar-01	32	4.0	0.4	-3.2	190	70	1.47	-0.5	213.68	2.56	0.00	0.04
05-Mar-01	36	4.0	0.5	-2.1	195	82	1.53	-0.5	189.26	2.54	0.11	0.13
09-Mar-01	40	4.0	0.4	-2.1	179	69	2.14	-0.2	315.23	2.99	0.23	0.06
13-Mar-01	44	3.0	0.4	-1.8	205	103	2.06	-0.3	203.55	3.91	0.04	0.46
17-Mar-01	48	3.0	0.3	-1.9	206	112	2.36	-0.2	214.33	4.08	0.04	0.20
21-Mar-01	52	3.0	0.3	-1.8	269	153	2.69	-0.3	178.80	4.26	-0.01	0.43
27-Mar-01	58	2.0	0.4	-1.9	243	139	3.13	0.0	228.82	4.20	-	0.19