# METEOROLOGICAL METRICS ASSOCIATED WITH DEEP SLAB AVALANCHES ON PERSISTENT WEAK LAYERS

by

Alex Grayson Marienthal

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Earth Sciences

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2014

©COPYRIGHT

by

Alex Grayson Marienthal

2014

# ACKNOWLEDGEMENTS

I would first like to acknowledge my parents, James and Emily, for raising me in the cold, snowy mountains, and for continuously supporting my passions while subtly enforcing good decisions and safety.

I owe gratitude to Jordy Hendrikx, Karl Birkeland, and Kathi Irvine for their academic guidance, and for being supportive through a few personal setbacks throughout this work. I also thank Mark Greenwood for statistical instruction and feedback on many analyses of my data. I thank all the professors and teachers that shared their knowledge and excitement for science and snow. These range from my teachers at Gold Hill elementary school, Steven Vanek and Pat McQuown, to professors Steve Custer and Stuart Challender at Montana State. From foundation to details, there are many teachers, friends, and family that have contributed to this work. Thank you.

Thanks to the Bridger Bowl ski patrol for sharing their thoughts on weather and avalanche forecasting, providing funding to attend and present at workshops, and collecting weather and avalanche data, and thanks to Ella Darham for thorough management of that data and making it available. Thanks to Scotty Savage, Mike Buotte, and the Big Sky ski patrol for their many discussions about deep slabs, and thanks to Randy Elliott and Doug Richmond at Bridger Bowl for many discussions about deep slabs, weak zone detecting x-ray goggles, and “the arc of the turn”.

Financial support for this work was provided by the American Avalanche Association Graduate Student Research Grant and a graduate teaching assistantship through the Earth Science Department at Montana State University.

TABLE OF CONTENTS

1. INTRODUCTION ..... 1

1.1 Background ..... 1
1.2 Research Questions ..... 4

2. LITERATURE REVIEW ..... 6

2.1 Snow Avalanches ..... 6
2.1.1 Slab Avalanche Fracture Mechanics ..... 8
2.1.2 Dry Slab Avalanche Formation and Initiation ..... 9
2.1.3 Wet Slab Avalanche Formation and Initiation ..... 15
2.2 Persistent Weak Layers ..... 17
2.3 Climax and Deep Slab Avalanches ..... 22
2.4 Applied Avalanche Forecasting ..... 31
2.5 Statistical Avalanche Forecasting ..... 36
2.5.1 Univariate and Multivariate Analysis ..... 37
2.5.2 Nearest Neighbors Analysis ..... 46
2.5.3 Classification Trees ..... 49
2.5.4 Random Forests ..... 58

3. METHODS ..... 60

3.1 Study Area ..... 60
3.2 Datasets ..... 64
3.2.1 Meteorological Observations ..... 64
3.2.2 Avalanche Records ..... 68
3.2.3 Snowpack Observations ..... 69
3.3 Imputations for Missing Meteorological Data ..... 70
3.3.1 Imputation Methods ..... 71
3.3.1.1 Temperature and Snow Depth Estimates ..... 71
3.3.1.2 Snow Depth Estimate Corrections ..... 72
3.3.1.3 New Snowfall and Snow Water Equivalent (SWE) Estimates ..... 73
3.3.2 Imputation Results ..... 75
3.3.2.1 Maximum Temperature ..... 78
3.3.2.2 Minimum Temperature ..... 79
3.3.2.3 Snow Depth ..... 79
3.3.2.4 New Snowfall and Snow Water Equivalent (SWE) ..... 81
3.3.3 Imputation Conclusion ..... 83
3.4 Historical Analysis of Avalanches at Bridger Bowl 1968-2013 ..... 83
3.4.1 Deep Slabs on Persistent Weak Layers ..... 84
3.4.2 Days and Seasons without Deep Slabs ..... 87
3.4.3 Meteorological Metrics ..... 88

TABLE OF CONTENTS – CONTINUED

3.4.3.1 Daily Meteorological Metrics ..... 89
3.4.3.2 Seasonal Meteorological Metrics ..... 90
3.4.4 Statistical Analysis Methods ..... 93

4. RESULTS ..... 100

4.1 Case Study: The 2011-2012 Season ..... 100
4.1.1 Daily Weather and Snowpack History ..... 101
4.1.2 Avalanche Activity ..... 105
4.1.3 Seasonal Meteorological Summary ..... 107
4.2 Historical Analysis: 1968/69-2012/13 ..... 110
4.2.1 Univariate Analysis Results ..... 110
4.2.2 Classification Trees for Seasons with Deep Slabs on Persistent Weak Layers ..... 112
4.2.3 Classification Trees for Days with Deep Slabs on Persistent Weak Layers ..... 117
4.2.4 Random Forests ..... 123

5. DISCUSSION ..... 125

5.1 Case Study: The 2011-2012 Season ..... 127
5.2 Historical Analysis: 1968/69-2012/13 ..... 133
5.2.1 Univariate analysis ..... 133
5.2.2 Classification Trees for Seasons with Deep Slabs on Persistent Weak Layers ..... 137
5.2.3 Classification Trees for Days with Deep Slabs on Persistent Weak Layers ..... 141
5.2.4 Random Forests ..... 146
5.2.5 Study Limitations ..... 148

6. CONCLUSION ..... 152

REFERENCES CITED ..... 157

APPENDICES ..... 167

APPENDIX A: Model Residuals for Imputations ..... 168
APPENDIX B: Snowpit Profiles From the 2011-2012 Season ..... 177
APPENDIX C: Classification Trees ..... 206

# LIST OF TABLES

Table Page

1. Contributory Factors to Avalanche Formation and Initiation ..... 10

2. Previous Studies that Examined Deep Slab Avalanches ..... 23

3. Previous Studies that Applied Statistical Methods for Avalanche Forecasting ..... 39

4. Meteorological Observations from the Alpine Weather Station that are Recorded Daily by Ski Patrol at Bridger Bowl ..... 66

5. List of Months that are Missing from Meteorological Records at the Alpine Weather Station ..... 67

6. Descriptive Statistics for Measured Meteorological Metrics at Each Weather Station ..... 67

7. Results of Model Selection for Weather Variable Imputations ..... 77

8. Corrections Applied to Estimated Snow Depth Values ..... 81

9. Summary of Snow Densities Recorded at Bridger Bowl, and in Various Locations of Other Studies ..... 83

10. Summary of the 24 Seasons that had at Least One Avalanche Failing on a Deep Persistent Weak Layer Late in the Season ..... 87

11. Daily Meteorological Metrics Derived to Compare Conditions that Precede Days with and Days without Deep Slabs on Persistent Weak Layers ..... 90

12. Seasonal Meteorological Metrics Derived to Compare Conditions During Seasons with and Seasons without Deep Slabs on Persistent Weak Layers ..... 91

13. Contingency Table Showing Definitions for Equations Used to Measure Classification Tree Performance ..... 98

14. Snowpack Observations During the 2011-2012 Season at Bridger Bowl ..... 103

LIST OF TABLES - CONTINUED

Table Page

15. Seasonal Meteorological Metrics for 2011-2012 Compared to Metrics for all 44 Seasons ..... 109

16. Summary statistics for All Seasonal Metrics and Results from the Wilcoxon Rank-sum Test ..... 111

17. Summary statistics for All Daily Metrics and Results from the Wilcoxon Rank-sum Test ..... 113

18. Contingency Table Values and Performance Measures for Classification Trees Grown to Split Seasons with and Seasons without Deep Slabs on Persistent Weak Layers ..... 114

19. Contingency Table Values and Performance Measures for Classification Trees Grown to Split Days with and Days without Deep Slabs on Persistent Weak Layers ..... 120

20. Measures of Model Performance for Cross-validated Trees that Split Days with Dry Deep Slabs, Days with Wet Deep Slabs, and Days without Deep Slabs ..... 123

21. Contingency Table for Random Forests ..... 123

22. Measures of Model Performance for Random Forests ..... 123

# LIST OF FIGURES

Figure Page

1. Avalanche Fatalities in the United States by Season 1950-2013 ..... 2

2. A Medium Sized Direct-action Slab Avalanche ..... 7

3. Crown of a Climax Slab Avalanche Near Bridger Bowl ..... 8

4. Depth Hoar Crystals that Formed at Bridger Bowl During the 2011-2012 Season ..... 18

5. Map Showing Location of Bridger Bowl Ski Area ..... 61

6. Annual Average Daily Skier Visits at Bridger Bowl Since 1985 ..... 62

7. Location of Weather Stations and Avalanche Areas at Bridger Bowl ..... 65

8. Time Series Plots of Modeled Alpine Snow Depth Values and Measured Values from Both Stations, Shown Before and After Applying a Correction to Modeled Values ..... 74

9. Plots of New Snowfall and New Snow Water Equivalent at the Bridger Bowl Weather Station Versus Those Metrics at the COOP Weather Station ..... 74

10. Plots of Temperature and Snow Depth Values at Bridger Bowl Versus the COOP Station ..... 76

11. Relationship Between New Snow Density and Maximum Temperature ..... 82

12. Overview of the Location and Order of Large Wet Slab Avalanches that Failed on Depth Hoar on March 27th, 2012 at Bridger Bowl ..... 101

13. Meteorological Observations from Bridger Bowl During the 2011-2012 Season ..... 102

14. ECT Results Plotted by Depth of Failure Layer during the 2011-2012 season at Bridger Bowl ..... 104

LIST OF FIGURES - CONTINUED

Figure Page

15. Daily avalanche activity, Snow Water Equivalent, and Temperature During the 2011-2012 Season at Bridger Bowl ..... 106

16. Over-fit Classification Tree Created from Seasonal Metrics Prior to January 1st to Split Seasons with and Seasons without Deep Slabs on Persistent Weak Layers ..... 114

17. Cross-validated Classification Tree Created from Seasonal Metrics Prior to January 1st to Split Seasons with and Seasons without Deep Slabs on Persistent Weak Layers ..... 115

18. Over-fit Classification Tree Using Seasonal Metrics Prior to February 2nd to Split Seasons with and Seasons without Deep Slabs on Persistent Weak Layers ..... 116

19. Cross-validated Classification Tree Using All Metrics Prior to February 2nd to Split Seasons with and Seasons without Deep Slabs on Persistent Weak Layers ..... 116

20. Over-fit Classification Tree Using All Seasonal Metrics to Split Seasons with and Seasons without Deep Slabs on Persistent Weak Layers ..... 117

21. Over-fit Classification Tree Grown Using Only Daily Meteorological Metrics to Split Days with and Days without Deep Slabs on Persistent Weak Layers ..... 118

22. Cross-validated Classification Tree Using Daily Metrics to Split Days with and Days without Deep Slabs on Persistent Weak Layers ..... 119

23. Cross-validated Classification Tree Grown to Split Days with and Days without Deep Slabs on Persistent Weak Layers Using Metrics that are Significantly Different Between Days with and Days without Deep Slabs ..... 120

24. Cross-validated Classification Tree Grown to Split Days with Dry Deep Slabs, Wet Deep slabs, or No Deep Slabs ..... 122

LIST OF FIGURES - CONTINUED

Figure Page

25. Cross-validated Classification Tree Grown to Split Days with Dry Deep Slabs, Wet Deep slabs, or No Deep Slabs Using Variables with Significant Differences Between Days with and Days without Deep Slabs ..... 122

26. Variable Importance Shown Using the MDA for Each Variable in the Random Forests for Daily Variables and Seasonal Variables ..... 124

# ABSTRACT

Snow avalanches are a potentially fatal and highly destructive natural hazard. Snow slab avalanches occur in steep alpine terrain due to an unstable layered snowpack. When a consolidated layer of snow forms a slab above a weak layer of snow the slab may collapse and slide downhill due to gravitational and applied forces (e.g., the weight of a skier, explosive, or new snowfall). Persistent weak layers form in the snowpack due to strong vapor pressure gradients, and they can last for weeks to months as a slab builds above them. Avalanches on persistent weak layers become less frequent, yet are typically larger and more destructive the longer and deeper the layer is buried. Deep slab avalanches on persistent weak layers pose a difficult forecasting problem due to their low likelihood of occurrence and potentially high consequences. This thesis aims to identify meteorological metrics that are associated with deep slabs on persistent weak layers. We used univariate analysis, classification trees, and random forests to explore relationships between seasons with deep slabs and summaries of meteorological metrics over the beginning of the season during weak layer formation. We also looked at the relationship between days with these avalanches and summaries of meteorological metrics over the days prior to them. In addition, we reviewed a case study of a season that had multiple deep slabs on a persistent weak layer and a historic wet slab avalanche cycle on the same layer, at Bridger Bowl ski area. Seasons with deep slabs typically had relatively low precipitation throughout the early part of the season (i.e., November - January), and a snowpack in the beginning of the season that was sufficiently deep, but shallow enough for a weak layer to develop. Our results also showed warmer twenty-four hour temperatures and more precipitation over seven day prior to days with dry deep slabs, and extended periods of above freezing temperatures were seen prior to days with deep wet slabs. These results are in line with previous research and are suggestive of meteorological summaries that may be useful to forecast deep slab avalanches on persistent weak layers.

# 1. INTRODUCTION

## 1.1 Background

Avalanches threaten to kill or injure people, and affect tourism, transportation and development, in alpine areas worldwide. Avalanches have been a familiar phenomenon in the European Alps and Scandinavia for more than five centuries, and have become a greater issue in other areas of the world over the past century (LaChapelle, 1960a). They were recognized as a problem in the western United States around 1890 during the gold rush. In 1910, 108 fatalities, three wrecked trains, and one million dollars in property damage resulted from a series of large avalanches on a railroad in the Northwestern United States (Atwater, 1954; Gallagher, 1967). The importance of understanding the avalanche phenomenon was further emphasized when approximately 10,000 U.S. troops were killed in a series of avalanches in Europe during World War I (Atwater, 1954). More recently, increased population and activities in alpine areas have resulted in an increase in human-avalanche encounters. In the United States, the average annual number of deaths from snow avalanches have historically exceeded that from earthquakes, and from all other forms of mass movement combined (Voight et al., 1990). The number of fatalities each season in the United States increased rapidly between 1990 and 2000 (Fig. 1). Since 2000, the rate of increase of fatalities per season has leveled off, but the total number remains high at around 30 per year (Fig. 1). The increase in fatalities can be largely attributed to better performance of snowmobiles and ski equipment, and their riders, which has placed many more in the way of avalanches (e.g., Chabot, 2002). Furthermore, activities that take place in avalanche terrain have gained popularity and the

number of people participating in them has increased. With more people facing encounters with avalanches there is an increasing need to better understand, and inform the public about, the processes that create and initiate avalanches.

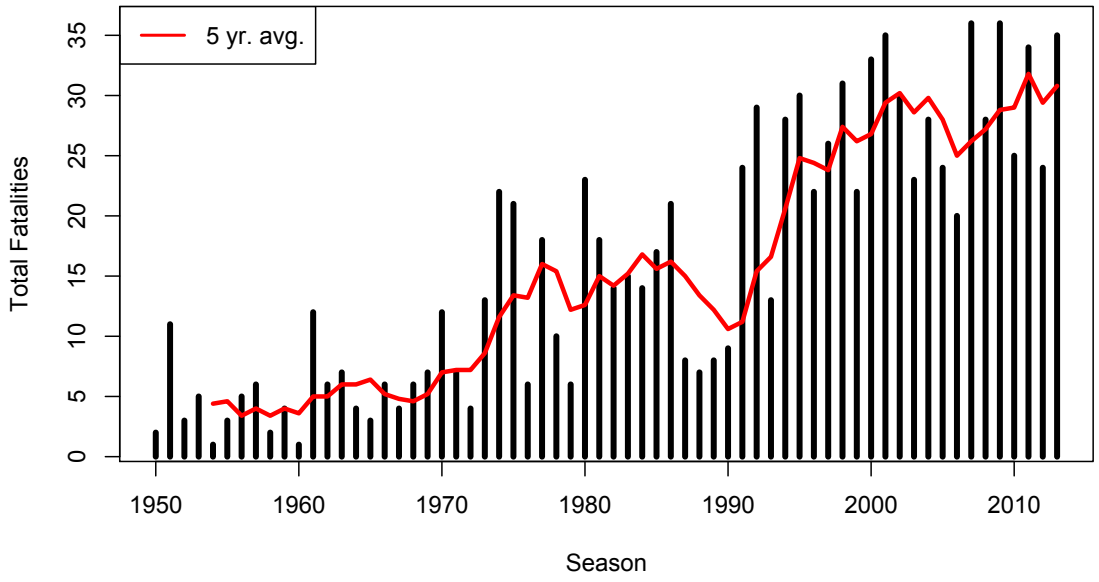



Figure 1. Avalanche fatalities in the United States by season 1950-2013 (Data: CAIC, 2014).

The general mechanisms of snow avalanches are relatively well understood (e.g., Schweizer et al., 2003). However, predicting the exact time and location of an avalanche is difficult due to the highly variable distribution of snow through space and time. The general location of avalanches can be predicted by aspect and elevation. The formation of the snowpack is highly dependent on exposure to the sky (i.e., radiation), sun, wind, precipitation, and temperature, which all have different effects dependent on aspect and elevation. Predicting more precise locations of avalanches, such as on a specific avalanche path, requires intimate knowledge of the snowpack and the snowpack history on that slope.

Avalanche danger over an area is derived from information about the general structure of the area's snowpack, signs of instability in the snowpack, and current weather conditions (LaChapelle, 1980). Knowledge of the snowpack structure is derived from digging snowpits and following the formation and development of layers throughout the season. Current weather conditions inform those forecasting avalanche hazard about how the snowpack is changing due to temperature changes or loading from wind and snow. Knowledge of the current weather conditions is more powerful when combined with information about the snowpack strength and structure. The state of the snowpack prior to changes in weather determines how the snowpack reacts to those changes.

Information about the snowpack strength is relatively obvious when the main avalanche problem, or weak layer, is fresh snow or a newly formed persistent weak layer. In these situations avalanche activity can be prevalent, and the young age (i.e., hours or days) of the layers indicates that they likely will gain strength over time. In addition to the relatively obvious instability of new layers, avalanches triggered on newer layers are often smaller and have lower consequences than avalanches triggered on old layers due to the greater depth of the slab overlying older layers. Weak layers that are buried and remain in the snowpack are called "persistent" weak layers. In general, a persistent weak layer gains strength over time (e.g., Logan et al., 2007) and the likelihood of triggering an avalanche on that layer goes down, but the likely size and consequence if an avalanche is triggered goes up. As a result, relatively few signs of snowpack weakness (i.e., avalanches, and collapsing snow) are observed prior to avalanches on old persistent layers. The lack of evidence available regarding the strength of the snowpack makes

prediction of avalanches on old persistent weak layers more difficult. With a relative lack of traditional indicators about the strength of the snowpack, alternative information may be helpful in forecasting avalanches on old persistent weak layers.

The meteorological conditions under which a weak layer forms, as well as the conditions subsequent to its formation, will influence its strength (e.g., Akitaya, 1974). Therefore, summaries of meteorological conditions during and following weak layer formation can be used to supplement other observations when forecasting for avalanches on a persistent weak layer that has become less active.

## 1.2 Research Questions

This thesis will explore relationships between meteorological conditions during the formation of persistent weak layers and avalanches that fail on these weak layers late in the season. Our focus is on avalanches events that were observed and recorded within and adjacent to Bridger Bowl ski area from 1968-2013. Deep slab avalanches, in this thesis, are assumed to have failed on a layer of depth hoar, or a layer of deeply buried facets (possibly combined with an ice crust). These types of layers commonly persist in the study area's climate and ski area setting. Depth hoar and basal facets are common persistent weak layers that form during the early months of winter (November-January in the Northern hemisphere) mainly in continental and intermountain climates. Avalanche activity on these layers can be abundant in December and January, and become less frequent going into February. Around this time less activity occurs on these deep layers, but when it does occur, larger avalanches are more common than smaller ones. This thesis defines avalanches that fail on deep persistent weak layers and focuses on those

believed to have failed on a deep persistent weak layer on or after February 1st at Bridger Bowl ski area in southwest, Montana.

The objective of this thesis is to improve the ability to forecast the timing of avalanches that fail on deep persistent weak layers. More specifically, the goal is to determine what additional data can be used to assess the instability of a weak layer once data that is commonly used to assess instability (i.e., avalanche occurrences or stability tests) gains more uncertainty as the weak layer ages. This will be achieved by answering two related questions: 1) what is the relationship between meteorological metrics during the early months of the season and deep slabs on persistent weak layers late in the season? Strong relationships here imply that forecasting avalanches on an old persistent weak layer begins with data gathered when the weak layer is formed, and that information must be retained through the season. 2) When a persistent weak layer is present, which metrics over the days prior to deep slabs on persistent weak layers commonly have different values than they do prior to days without deep slabs? This question will address what current changes in weather typically contribute to the initiation of deep slabs on persistent weak layers. In addition to statistical analysis of historical data, a case study of the 2011-2012 season at Bridger Bowl will provide insight and context to these questions. The 2011-2012 season had a layer of depth hoar that produced avalanches throughout the entire season and culminated with an historic wet slab cycle that took out the entire season's snowpack.

# 2. LITERATURE REVIEW

## 2.1 Snow Avalanches

There are two general types of avalanches: loose-snow avalanches and slab avalanches. Loose snow avalanches are the result of mostly unconsolidated snow grains at the snow surface that begin to slide from a point and fan out downslope as more and more grains are sent into motion through collisions with each other (McClung and Schaerer, 2006). Slab avalanches are the result of one or more layers of well-bonded snow releasing simultaneously from the snow cover and sliding down a slope due to poor bonding between the sliding layers and the layers of snow, or ground (Fig. 2) (Schweizer et al., 2003). Slab avalanches are usually the most dangerous and destructive type of avalanche, and are the primary target for mitigation or avoidance (LaChappelle, 1960a). The focus of this thesis is slab avalanches.

Both types of avalanches are further sub-divided into wet and dry snow avalanches. This is based on whether the snow is wet or dry where the avalanche initiates (i.e., in the starting zone). Dry snow avalanches are faster moving and are usually accompanied by a fast moving powder cloud made up of snow crystals and air. Wet snow avalanches are slow moving and can be highly destructive due to their high density. Intermediate types of avalanches exist, being composed of both dry and wet snow (LaChappelle, 1960a; McClung and Schaerer, 2006).



Figure 2. A medium sized (SS-AS-R3-D2) direct-action slab avalanche in Southwest Montana. This avalanche was initiated on low angle terrain behind (downhill from) where the photo was taken. The fracture propagated to the steeper slope to the right of the photo, and then propagated up to the less steep slope above. The arrested pieces of slab indicate a lower slope angle.

Slab avalanches occur as either a direct action avalanche or a climax avalanche (LaChapelle, 1960a; Schweizer et al., 2003). Direct action avalanches are the immediate result of a single storm or meteorological event (Fig. 2). Climax avalanches are not necessarily associated with a single storm or meteorological situation, and are often avalanches involving snow layers deeper than the previous storm (Fig. 3) (LaChapelle, 1960a). This thesis involves both dry and wet, climax slab avalanches.

To most effectively present this research, a necessary background on snow avalanche formation and initiation will first be established. Slab avalanche fracture mechanics will be discussed briefly, followed by a discussion of factors contributing to dry and wet slab avalanche formation and initiation. This will be followed by a

discussion of the formation and properties of common persistent weak layers, forecasting avalanches and climax avalanches, and statistical avalanche forecasting.



Figure 3. Crown (.3-2m deep) of a climax slab avalanche (HS-N-R5-D4) near Bridger Bowl. This avalanche released naturally on March 20, 2012 on a 30-60cm thick layer of depth hoar that formed early in the season.

### 2.1.1 Slab Avalanche Fracture Mechanics

Slab avalanches result from the propagation of a fracture beneath and along all lateral boundaries of a slab (McClung and Schaerer, 2006). Initiation of a slab avalanche is due to rate-dependent changes to the snowpack. These occur from either rapid load from a localized trigger near the surface (i.e., a skier or explosive), widespread gradual loading (i.e., from precipitation or wind), or, in the absence of load, temperature or radiation induced changes to the snow cover's properties (e.g., warming of the surface) (Schweizer et al., 2003). These triggers were previously thought to initiate a fracture in the weak layer below a slab through propagation of shear slip surfaces. This theory generally states that differential rates of deformation between the slab and the weak layer

lead to shear slip faces at the interface that rapidly propagate when a critical length of the slip face is reached (McClung, 1979; McClung, 1981; McClung, 2005).

The general theory of shear stress overcoming shear strength to initiate an avalanche is useful to explain why a steep slope is required for an avalanche to slide. However, the frequently observed triggering of avalanches and fracture of weak layers in low angle terrain, often accompanied by a "whumph" sound, is unexplained by this shear fracture mechanism (Heierli et al., 2008). While previous work suggests collapse of the slab to occur entirely after shear failure (McClung, 2005), recent work suggests a combination of slope-normal failure (collapse) and slope-parallel failure (shear) during avalanche release (Heierli and Zaiser, 2008). More recent studies have shown that frictional forces along the fracture plane, directly after collapse and de-bonding of the slab from the other layers, must be overcome by gravitational forces on a steep slope (van Herwijnen and Heierli, 2009; Heierli et al., 2011). This suggests collapse is the primary process in avalanche initiation followed by the forces of friction being overcome by forces from normal load (van Herwijnen and Heierli, 2009; Heierli et al., 2011).

### 2.1.2 Dry Slab Avalanche

#### Formation and Initiation

Avalanche formation is a result of the interaction between the terrain, existing snow cover, and meteorological conditions (Atwater, 1954; Schweizer et al., 2003; McClung and Schearer, 2006). Atwater (1954) recognizes the need for a slope (terrain) and snow to create an avalanche. He mentions ten snow related factors that contribute to avalanche formation and initiation (Table 1). Schweizer et al. (2003) refine these ten factors to terrain plus four snow related factors that contribute to avalanche formation,

which are: new snowfall, wind, temperature and radiation, and snow cover stratigraphy (Table 1).

Table 1 Ten factors that contribute to avalanche initiation and formation from Atwater (1954), and four refined factors from Schweizer et al. (2003).

| | Atwater, 1954 | Schweizer et al., 2003 |
|---|---|---|
| 1 | Depth of snow on the ground | Snow cover stratigraphy |
| 2 | Character of the old snow surface | |
| 3 | Settlement | |
| 4 | New snow depth | New snowfall (amount, rate, and density) |
| 5 | New snow type | |
| 6 | Amount of water in new snow | |
| 7 | Snowfall intensity | |
| 8 | Precipitation intensity (Amount of water/time) | |
| 9 | Wind action (Speed and direction) | Wind |
| 10 | Temperature | Temperature and radiation |

Terrain is a factor that essentially remains constant throughout avalanche formation. Terrain influences can be slightly altered by the spatial variability of snow accumulation. In deeper snow climates bed surface roughness can be reduced, and slope angles can increase or decrease with more accumulation. The primary terrain factor that contributes to avalanches is the requisite slope steep enough to allow snow to slide after failure (Atwater, 1954; McClung and Schaerer, 2006). Slab avalanches are usually initiated on slopes between 30° and 45° (LaChapelle, 1960a; Perla, 1977; Schweizer and Jamieson, 2001; Schweizer et al., 2003). Micro-topography in the terrain also influences the formation and initiation of avalanches. Slope-scale concavities contribute to avalanche formation by being conducive to the accumulation of more snow (Gleason, 1995; Schweizer et al., 2003). Convexities in the terrain promote areas of stress on the

snowpack, which can contribute to avalanche initiation (Föhn et al., 2002; Schweizer et al., 2003; McClung and Schaerer, 2006).

Terrain roughness also plays a role in avalanche formation and initiation. Small objects such as rocks and shrubs can interrupt the formation of continuous layers, and act as anchors or create areas of stress on slabs when the snow is shallow and they protrude above the surface (Schweizer et al., 2003; McClung and Schaerer, 2006). When the snowpack becomes deeper, the terrain roughness is smoothed out, and may allow for the formation of more continuous weak layers and slabs. Thus, snow depth prior to storms is an important contributing factor to the formation of avalanches (Schweizer et al., 2003; Schweizer et al., 2009). More variability in terrain roughness can create isolated areas of weak faceted crystals where the snow cover is thin and the temperature gradient is strong around shallow objects (Arons et al., 1998; Schweizer et al., 2003). These can become areas where triggering an avalanche is more likely (McClung and Schaerer, 2006).

Vegetation can also interrupt the formation of continuous layers and reduce loading. Dense forests intercept precipitation and snow transported by wind, and alter the radiation transfer (Schweizer et al., 2003; McClung and Schaerer, 2006). Terrain and vegetation interact with the existing snow stratigraphy and meteorological conditions to create avalanches.

In addition to needing terrain that can produce an avalanche, the structure of the snowpack must promote the formation of an avalanche. In order to have an avalanche, the snow stratigraphy must contain a buried weak layer underlying a slab (Schweizer et al., 2003; McClung and Schaerer, 2006). Schweizer and Jamieson (2001) found that 70 percent of 186 skier-triggered avalanches had a weak layer composed of a persistent

grain form, such as surface hoar, near-surface facets, or depth hoar (Schweizer et al., 2003). Therefore, avalanches on persistent types of weak layers deserve attention. The properties of the slab above, and to a lesser extent below, the weak layer must be considered in addition to the weak layer. Slab depth, hardness, moisture content, temperature, and grain type can affect the potential for fracture propagation (McClung and Schaerer, 2006).

An avalanche will initiate on a steep slope in a snowpack with stratigraphy conducive to failure when the meteorological conditions promote changes in the snowpack that lead to a fracture, or collapse, of the weak layer. New precipitation has been regarded as the strongest parameter to forecast avalanches (e.g., Bois et al., 1975; Föhn et al., 1977; Schweizer and Föhn, 1996; Davis et al., 1999; Floyer and McClung 2003; McCollister et al., 2003). Amount, intensity, and characteristics of new snow contribute to avalanche formation and initiation (e.g., Perla 1970; Judson and Erickson, 1973; Bovis, 1977). Wind is often regarded as the second most significant contributing factor to avalanche formation because it has a similar effect as new snow (Atwater, 1954; Schweizer et al., 2003). Snow transported by wind adds load and creates slabs. Wind can be difficult to measure due to high variability in local wind speed and direction. This high variability also contributes to the spatial variability of the snow stratigraphy, and stress concentrations within the snowpack, by creating slabs of spatially heterogeneous hardness and thickness (Schweizer et al., 2003). Parameters that combine wind and snowfall have also been found as strong indicators of avalanche activity. These parameters often represent wind drifting, or loading from wind, by combining snow

available for transport (new snow) with wind speed and direction (e.g., Schweizer and Föhn, 1996; Davis et al., 1999; Hendrikx et al., 2005).

While the addition of weight by new snowfall, rain, or blowing snow is the most common natural trigger for dry slab avalanches, the rate of loading from these plays an important role in avalanche initiation. Conway and Wilbour (1999) suggest that during storms, there is a balance between the rate at which stress is added to the snowpack and the rate at which it can adjust its strength to support the addition of load. If the snowpack cannot adjust its strength as fast or faster than the rate of loading, an avalanche will result.

Temperature can be a contributing factor in dry slab avalanche initiation, specifically when there is an absence of new load (Schweizer et al., 2003). Although temperature has received considerable attention in literature (e.g., Schweizer et al., 2003) and educational material (e.g., McClung and Schaerer, 2006) it is believed to have a relatively minor effect on dry slab avalanche initiation (McClung, 1996; Schweizer and Jamieson, 2010). For dry slab avalanches, temperature primarily affects surface layers due to the low thermal conductivity of low-density snow (Sturm et al., 1997; Schweizer et al., 2003). The snowpack is generally warmer, denser, and stronger deeper in the snowpack (McClung, 1996). As with most meteorological effects on the snowpack, the rate of change is important. Rapid temperature change can lead to increased slab deformation, which is accompanied by increased strain and strain rates at the slab-weak layer interface (Schweizer et al., 2003). Previous work and observations suggest that rising slab or weak layer temperatures, which result from rising air temperatures, can lead to decreased stability (McClung, 1996). The immediate effects of warming typically

decrease stability, while delayed effects of warming typically increase stability (McClung and Schweizer, 1996). The delayed effects of warming promote stability through increased bond formation, reduced temperature gradients, and increased creep rates, which increase settlement, densification, and hardness. Immediately following warming, reduced slab hardness is the primary factor that results in decreased stability. Furthermore, this reduction in hardness makes skier triggering more likely as skiers will affect softer snow to a greater depth than harder snow (McClung and Schweizer, 1996; Schweizer et al., 2003). Surface warming is suggested to be more relevant in skier triggered avalanches than in natural releases (Harvey and Heierli, 2009; Schweizer and Jamieson, 2010).

Surface warming makes skier triggering more likely, and increases shear stress in the weak layer, even if the warming does not reach all the way through the slab to the weak layer (Wilson et al., 1999). Schweizer and Jamieson (2010) suggest that dry slab avalanche initiation due to warming is clearly not due to warming of the weak layer, leading to decreased strength. They mention that the effect of surface warming on stability is strongly dependent on the stratigraphy, and on the slab properties. This suggests that surface warming is most effective when the slab is a relatively thin new snow layer. Despite empirical evidence of deep slabs during warming (e.g., Schweizer et al., 2009; Conlan et al., 2014) there is not an explainable causal effect for the coincidence (Schweizer and Jamieson, 2010).

Radiation has an equal or greater influence than temperature on surface warming that contributes to avalanche formation and initiation (Schweizer et al., 2003; Schweizer and Jamieson, 2010). Radiation data have been applied to avalanche forecasting, with

varying success, mainly in regards to wet slab avalanches (Baggi and Schweizer, 2009; Peitzsch et al., 2012). For dry slab avalanches, the result of surface warming by radiation has results similar to warming by temperature (Schweizer and Jamieson, 2010). Long-wave radiation being emitted from the snow surface during clear days and nights can counteract warming from direct shortwave (solar) radiation (Schweizer et al., 2003). An opposite effect occurs on cloudy days when long-wave radiation is emitted from the clouds back to the snow, which enhances warming of the snowpack. This can lead to wet snow avalanches (Tremper, 2008; McClung and Schaerer, 2006).

### 2.1.3 Wet Slab Avalanche

#### Formation and Initiation

The factors contributing to slab avalanche formation and initiation vary between wet and dry slab avalanches. In general, the initiation of a dry slab avalanche is the result of stress being added to the snowpack quicker than the snowpack can adjust its strength, while wet slab avalanche initiation is the result of a decrease in strength of the snowpack that allows it to succumb to existing, and added, stresses (Tremper, 2008). While these general processes dominate the respective types of avalanche initiation, neither one is exclusive of the other (McClung and Schaerer, 2006). It is likely that a decrease in strength or alteration of snowpack properties often accompanies the addition of stress during dry slab avalanche initiation. It is also common for wet slab avalanche initiation to be the result of an addition of stress, in conjunction with a decrease in snowpack strength (e.g., Baggi and Schweizer, 2009; Marienthal et al., 2012). The factors discussed in the previous section were primarily presented in the context of dry slab avalanches. Factors primarily contributing to wet slab avalanche initiation will follow.

The addition of free water to the snowpack is a primary contributing factor to the initiation of wet slabs (LaChapelle, 1960a; Kattelmann, 1984; Conway and Raymond, 1993; Reardon and Lundy, 2004; Baggi and Schweizer, 2009; Peitzsch et al., 2012). Measurements of SWE loss or snow settlement, and sustained warming, which suggest the introduction of water to the snowpack, have been used to forecast wet slabs (Baggi and Schweizer, 2009; Peitzsch et al., 2012). For wet slabs to occur, a similar stratigraphy is required as for dry slabs, though the processes leading to failure may differ. Weak layers, crusts, ice lenses, or capillary barriers (a significant difference in grain size between adjacent layers), that may impede water flow and allow water to flow along the layer boundary below a slab rather than through the snowpack, contribute to wet slab formation. Baggi and Schweizer (2009) found the presence of capillary barriers, increased load on a weakened snowpack, and days since the snowpack went isothermal to be useful for forecasting wet slabs in Davos, Switzerland.

Reardon and Lundy (2004) describe a structure for wet slab avalanches that includes a weak basal layer. While non-basal weak layers have also been observed as failure planes for wet slabs upon the interaction with free water (e.g., Conway and Raymond, 1993), they are less frequently an issue in ski area settings due to the frequent disturbance of the snowpack (Kattelmann, 1984). Wet slab activity was observed to peak during the transition from winter to summer snow (isothermal snow) in Glacier National Park, Montana (Reardon and Lundy, 2004), which suggests wet slab activity to occur on days near the time the snowpack turns isothermal. LaChapelle (1960a) also noted observing the most extensive wet slab cycles during the first warm up, or during winter rain.

Rain on snow events along the Milford Road in New Zealand were found to increase avalanche danger until vertical drainage channels were established through the snowpack, at which point avalanche danger would become low (Conway et al., 2009). Conway et al. (2009) identified three stages of water infiltration through snow following the start of rain on snow. These are generally: 1) immediately increased avalanche danger, 2) Danger of deeper avalanches prior to established drainage channels, and 3) a return to stability once drainage is established. Conway and Raymond (1993) observed the majority of wet slabs to occur in the upper 30-50cm of the snowpack directly after the onset of rain (1-2hr), except when the snowpack was relatively stable before. However, they noted that avalanches that had a delayed response (up to 13 hours) to rain were deeper and harder to predict.

The potential for wet slabs on basal layers in a ski area setting combined with their hard to predict nature suggests a need to better understand the factors that contribute to them. Although the variables described as contributing to wet slabs in these studies are limited to their respective study areas, they may be indicative of the types of variables to examine elsewhere. This thesis considers many of the aforementioned contributing factors to both wet and dry slab avalanches as contributing factors to climax wet and dry slab avalanches that fail on persistent weak layers.

## 2.2 Persistent Weak Layers

Depth hoar is a large angular snow crystal that makes up a common persistent weak layer at the base of the snowpack (Fig. 4). Depth hoar was found to account for 13% of fatal avalanche accidents in Canada between 1972 and 1991 (Jamieson and

Johnston, 1992). In 45 human-triggered avalanches in Switzerland that had accompanying snow profiles, approximately 20% failed on top of or within a layer of depth hoar (Schweizer and Lütschg, 2001). Depth hoar made up approximately 10% of weak layers associated with 186 investigated skier-triggered avalanches in Canada and Switzerland between 1987 and 1999 (Schweizer and Jamieson, 2001). Föhn (1993) found depth hoar to be the weak layer in 15% of roughly 300 snowpit profiles that were dug on a slope before or after it avalanched. In southwest Montana, depth hoar was the weak layer in 6% of 51 large backcountry avalanches investigated between 1990 and 1996 (Birkeland, 1998). Tracz (2012) found depth hoar to be the weak layer grain type in 60% of deep slab avalanches (Defined as avalanches with a persistent weak layer older than 3 weeks, and a crown depth exceeding the 80$^{th}$ percentile of crown depths in the region) in the Rocky Mountain region of Canada, while it made up the weak layer in an insignificant percentage of deep slabs in other regions of western Canada.

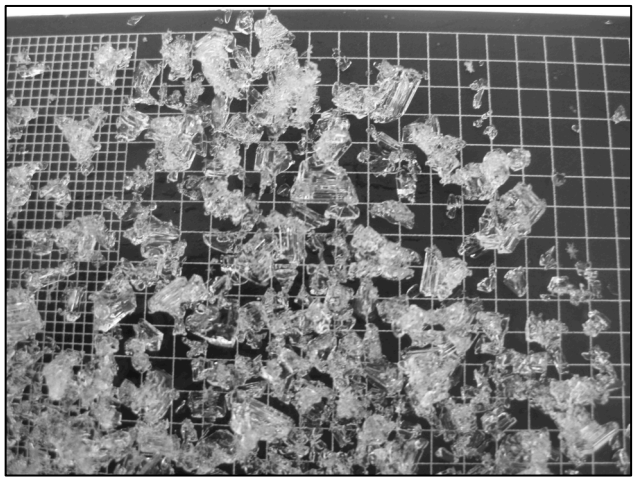


Figure 4. Depth hoar crystals that formed at Bridger Bowl during the 2011-2012 season. Grid squares are 1x1mm on the left and 3x3mm on the right.

Depth hoar is a common persistent weak layer in continental climates, where snow depth is shallow and temperatures are cold (Mock and Birkeland, 2000; Tremper, 2008). It forms from constructive metamorphism induced by rapid vapor transfer due to a strong vapor pressure gradient through the snowpack (Giddings and LaChapelle, 1962; Akitaya, 1974). This vapor pressure gradient drives water vapor from the base of the snowpack towards the top. Crystals grow larger from deposition of ice on their base during the upward movement of water vapor. The vapor pressure gradient can be observed by measuring the temperature gradient through the snowpack. Constructive metamorphism occurs when the vapor pressure gradient through the entire snowpack is stronger than the pressure gradient around individual grains, which is typically when the temperature gradient exceeds 10°C/m through the snowpack (LaChapelle, 1960b; Giddings and LaChapelle, 1962; Akitaya, 1974, Armstrong, 1985). A negative temperature gradient is formed in alpine snow covers due to below freezing snow surface temperatures and a ground temperature near 0°C, which is a result of geothermal heating and stored summer heat (McClung and Schaerer, 2006). In the absence of a strong negative temperature gradient, destructive metamorphism occurs (LaChapelle, 1960b). Destructive metamorphism is the movement of water vapor from points of high vapor pressure to low vapor pressure, which are due to irregularities in curvature, across the surface of individual grains. This process results in rounded grain forms, as opposed to the angular grains that form from constructive metamorphism (LaChapelle, 1960b; Giddings and LaChapelle, 1962).

Other factors that influence the development of depth hoar are snow temperature, initial crystal type, and initial density (Akitaya, 1974; Marbouty, 1980; Pfeffer and

Mrugala, 2002). Initial crystal type and density are highly correlated and also determine porosity, which plays a large role in vapor transfer and heat conduction during snow metamorphism (Bradley et al., 1977; Marbouty, 1980; Pfeffer and Mrugala, 2002). Initial snow densities greater than 350kg/$m^3$ have been shown to metamorphose into hard depth hoar with hardness ICSI R3 or greater (Colbeck, 1990) when exposed to strong temperature gradients (Akitaya, 1974; Marbouty, 1980; Pfeffer and Mrugala, 2002). Marbouty (1980) observed recrystallization in higher initial snow densities, but crystal growth was restricted by pore space. Similarly, snow with an initial density of 400kg/$m^3$ was found to metamorphose into hard (ICSI R3 or greater) solid type depth hoar (facets) connected by necks when exposed to temperature gradients between 20°C/m and 80°C/m (Pfeffer and Mrugala, 2002). Depth hoar resulting from the exposure of high density snow to a strong temperature gradient exhibits greater strength, implied by hardness, than depth hoar developed from lower density snow that has sufficient pore space for crystal growth (Akitaya, 1974; Pfeffer and Mrugala, 2002). When large depth hoar grains form from low-density snow there are much fewer bonds per unit area, resulting in a weak layer of snow (Akitaya, 1974; Pfeffer and Mrugala, 2002).

Average crystal size of depth hoar has been found to decrease with the snow temperature during formation (Akitaya, 1974; Marbouty, 1980). This typically produces larger crystals at the base of the snowpack since temperature generally decreases from the bottom to the top of the pack (Bradley et al., 1977; Marbouty, 1980). The main factor responsible for larger crystal sizes forming at higher temperatures is the amount of water vapor available for transfer and deposition (Akitaya, 1974) since more water vapor can be held in warmer air (Wallace and Hobbs, 2006 p. 64).

Bradley et al. (1977) found large fully metamorphosed depth hoar at the base of the snowpack to be relatively stronger than slightly less developed depth hoar above it. They identified the less developed depth hoar as the weakest zone, having forms between fully developed crystals with cups and striations and flat angular plate-like crystals. They also found the strongest temperature gradient to be in this weaker zone with less developed depth hoar. This stronger temperature gradient was explained by the lower conductivity of smaller bonds in the less developed depth hoar having to accommodate a large heat flux from the larger crystals below (Bradley et al., 1977).

Although depth hoar is the weak layer for a relatively lower percentage of avalanches than other persistent weak layer types, it is one of the more prominent persistent weak layer types in many ski area settings (e.g., LaChappelle and Atwater, 1961; Kattelmann, 1984; Marienthal et al., 2012). This is due to being more resistant to frequent disturbance than thinner weak layer types that form higher in the snowpack, such as surface hoar (LaChapelle and Atwater, 1961; Kattelmann, 1984).

At ski areas where depth hoar is a common weak layer, extensive efforts are made early in the season to reduce the formation and initiation of avalanches on depth hoar (e.g., Sahn, 2010). Explosives are used in a variety of ways to mitigate avalanches on depth hoar. This includes, but is not limited to, widespread application of explosives before depth hoar is deeply buried in order to disrupt the continuity of the layer (e.g., Carvelli, 2008). When this is not entirely successful, explosives are used to trigger avalanches on the depth hoar layer, ideally before they become relatively large and destructive. Typically, the later in the season that an avalanche on depth hoar is triggered, the larger it is. In addition to explosive application, compaction of depth hoar by boot

packing and skiers before it is deeply buried can effectively break up and strengthen the layer (LaChapelle and Atwater, 1961; Carvelli, 2008; Wieland et al., 2012).

Direct measurements of snowpack temperature gradients, snow density, and grain type are ideal for forecasting avalanches on persistent weak layers like depth hoar. However, these data are sparse in many cases even where meteorological data exist more thoroughly (e.g., a ski area). Meteorological metrics associated with the factors that influence the development of persistent weak layers in continental climates will be used in this thesis to proxy the factors that influence the formation and subsequent strength of these persistent weak layers. Air temperature, snow depth, daily new snowfall, and daily new snow water equivalent (SWE) will be used to derive variables that summarize the meteorological conditions in November, December, and January when persistent weak layers form, which may later cause deep slab avalanches.

## 2.3 Climax and Deep Slab Avalanches

Various studies have focused on climax avalanches and deep slab avalanches (Table 2). They seek to identify weather and snowpack variables that are most associated with these types of avalanches. While the exact definition of "deep slab" avalanches varies throughout these studies (Table 2), they all share a highly destructive and difficult to forecast character. The various definitions are an artifact of different climate types and available datasets.

Table 2 Characteristics and findings from previous studies that examined deep slab avalanches.

| Reference | Location | Type(s) of avalanche studied | Significant Findings |
|---|---|---|---|
| LaChapelle and Atwater (1961) | Alta, Utah and Squaw Valley, California | Climax avalanches | Dry early season prior to deep slabs on depth hoar, little warning immediately preceding deep slabs |
| Bradley and Bowles (1967), and Bradley (1970) | Bridger Bowl, Montana | Climax avalanches that fail on depth hoar | Shady (North) aspects have avalanches earlier in the season, Sunnier aspects have avalanches later |
| Fitzharris and Schaerer (1980) | Roger's Pass, Canada | Seasonal summaries of all avalanches | Characterized two climatological patterns common to winters that have more, or larger, avalanches |
| Jamieson et al. (2001) | Columbia Mountains, Canada | Avalanches failing on facets above a crust | Cumulative precipitation, air temperature change over 4-5 days, a shear frame stability index, a large hardness difference between the slab and weak layer. |
| Savage (2006) | Big Sky, Montana | Avalanches with crown depth >1.2m | Weak correlations with 24-48hr preceeding precipitations. Small explosives most common trigger |
| Comey and McCollister (2008) | Jackson Hole, Wyoming | Avalanches with crown depth > 1.18m | Small explosives most common trigger |
| Schweizer et al. (2009) | Davos, Switzerland | Large avalanches in one path | Substantial snow loading, snow depth deeper than the terrain roughness, a slight increase in air temperature, and a snowpack with a weak base |
| Tracz (2012) | Western Canada | Crown depth > 80th percentile of all crown depths in a region or > 0.8m | Precipitation over preceeding days, changes in air temperature and hours that air temperature is above freezing |
| Conlan et al. (2014) | Western Canada | Avalanches failing after the initial cycle of the instability on which they occur | Precipitation and positive temperature changes |
| Marienthal (2014) | Bridger Bowl, Montana | Days and seasons with deep slabs (>0.9m) on persistent weak layers after Feb. 1 | Dry early season with snow depth conducive to weak layer development, high cumulative precip. prior to days with deep slabs, sustained warming prior to deep wet slabs |

LaChapelle and Atwater (1961) mention delayed action climax avalanches as one of the most difficult forecasting problems. Many slopes reach the critical instability for a delayed action avalanche, but don't fail due to a lack of trigger, and then gain strength again. This can lead to a higher potential for delayed action climax avalanches in a ski area setting due to the numerous triggers (LaChapelle and Atwater, 1961). The best time to see the weak layers responsible for these events may be before they are buried (LaChapelle and Atwater, 1961). If this opportunity is missed, then the weak layer must be observed in a snowpit, or the meteorological conditions during weak layer formation may be useful to provide evidence of its strength or weakness.

Forest Service case histories from USFS snow rangers were used to describe seasons that had deeply buried weak layers, which eventually caused climax avalanche hazard. Three seasons (two at Alta, Utah and one at Squaw Valley, California) were described storm by storm from the USFS records (LaChapelle and Atwater, 1961). One season at Alta and one at Squaw Valley were characterized by a buried ice crust. The third season, at Alta beginning in the fall of 1959, was characterized by a buried layer of depth hoar. This layer produced erratic patterns of avalanche hazard, and then culminated with an extraordinary single climax avalanche (LaChapelle and Atwater, 1961). The snowfall in November and December of 1959 barely sufficed as a base for skiing. Snowfall in late December produced highly sensitive slabs on the depth hoar. Storms accompanied by wind in early January produced wind slabs failing on the depth hoar. This was followed by no significant storms until February 2-3, when a 64cm storm total lead to fewer than expected climax avalanches. On February 8-10 an 84cm storm total

lead to a climax avalanche that was the largest observed on that path over the preceding twenty years (LaChapelle and Atwater, 1961).

Further notable characteristics of the 1959 depth hoar season provide insight into forecasting and mitigating depth hoar hazard late into the season. Often, few signs of instability (e.g., natural activity, "whumphing", artificial activity) existed in paths that eventually produced climax avalanches, even with repeated application of explosives, until a critical load on the weak layer was reached (LaChapelle and Atwater, 1961). The absence of avalanches on depth hoar on north facing slopes in February was noted. This may suggest that more developed depth hoar on shady slopes can easily get cleared out with early storms and mitigation. Shady slopes hold more snow early season, so the first layers of depth hoar are preserved there while they are often destroyed on sunny slopes (Bradley, 1970). In addition, shady slopes typically harbor a cooler snow surface that is conducive to a stronger temperature gradient. This creates weaker layers of depth hoar that fail early while depth hoar on other aspects withstands the overburden pressure and begins a balancing act between strength and weakness (Bradley, 1970). The depth hoar layer at Alta in 1959 was noted to produce activity primarily on shaded slopes, while the ice-crusts in other seasons produced activity on all aspects (LaChapelle and Atwater, 1961). Furthermore, boot packing in three slide paths proved to be successful, as no avalanche activity was observed on the depth hoar in these paths. LaChapelle and Atwater (1961) plainly state, "Early-season depth hoar in a shallow snow cover is a danger sign which cannot be ignored."

Bradley and Bowles (1967) and Bradley (1970) studied deep climax slab avalanches that failed on basal depth hoar layers. They described a cyclical strengthening

and weakening of the weak layer through the season and attributed it to the balance between weakening from constructive metamorphism and strengthening from consolidation due to overburden pressures. These processes are primarily a result of loading from precipitation and strong temperature gradients through the snowpack during the months preceding climax slab avalanche activity. Depth hoar on north facing and more shaded slopes failed mostly in January, while depth hoar on sunny south facing slopes failed in March and April (Bradley, 1970).

Fitzharris and Schaerer (1980) analyzed climatological factors related to major avalanche winters affecting a stretch of the Canadian Pacific Railway across Roger's Pass in Canada. They used three avalanche summaries for each winter: number of avalanches to hit the railway, total estimated mass of those avalanches, and total length of railway covered by those avalanches. There was a weak correlation between these daily avalanche summaries and peaks in winter snowfall. Two common climatological patterns associated with major avalanche winters were identified. The first was characterized by a strong zonal flow, frequent storms, rapid temperature changes, and sustained snowfall. The other climatological pattern associated with major avalanche winters was characterized by a strong meridional flow, a sustained thermal regime that rapidly broke down or changed to a different regime, unusually extreme temperatures, and sustained intense presence of Arctic air that was followed by a major storm with temperatures rapidly rising to above 0°C and heavy precipitation. These major avalanche winters are more typical of those that have catastrophic climax slab avalanches on basal weak layers (i.e., the topic of this thesis), while the former type of major avalanche winter typically has many frequent, but small magnitude, avalanches (Fitzharris and Schaerer, 1980).

Jamieson et al. (2001) studied weather and snowpack factors associated with the release of avalanches on old deeply buried layers of facets that formed on crusts during two winters in the Columbia Mountains of western Canada. They found avalanches on these old buried facets to be associated with precipitation over several days preceding the avalanche, air temperature changes over 4-5 days, a shear frame stability index (Perla, 1970), and a large hardness difference between the slab and weak layer. A positive change in air temperature over the preceding 4-5 days showed a strong correlation with the avalanche activity on the facet layer. This suggests that warming and temperature effects play a role in the release of deep slab avalanches (Jamieson et al., 2001). The meteorological conditions during the formation of the facet layer were summarized from both the North Columbia Mountains, where the layer was well developed and produced many avalanches, and the South Columbia Mountains, where no activity was observed on this layer. Summaries of temperature, snow depth, and rain were used to convey the difference in meteorological conditions (i.e., temperature gradient based on air temperature, snow depth, and liquid water content) that allowed for the contrast in behavior of the facet layer between the two regions (Jamieson et al., 2001). This work is supportive of the use of meteorological summaries during weak layer formation to represent the characteristics of a weak layer later in the season.

A study by Savage (2006) examined avalanches with crown depths greater than 1.2m at Big Sky, Montana. Correlations with these avalanches were stronger for precipitation over the preceding three to five days than for precipitation over the preceding 12-36 hours (Savage, 2006). Weak interfaces were mostly thin facet layers. Depth hoar was not an associated weak layer in this dataset because avalanches on depth

hoar frequently slide earlier in the season with slab depths less than 0.8m (Savage, 2006). In addition to weak correlations with 12-36 hour preceding precipitation, 38% of these deep avalanches were on days without new snow and 82% that occurred with new snow were on days with less than 10cm of new snow. This is likely due to the fact that large avalanches are often triggered after strong winds at Big Sky. Furthermore, most deep avalanches were triggered with explosives, and the author notes the majority of explosive triggers were relatively small (1kg) hand charges (Savage, 2006). The small size of explosives suggests that the snowpack is near critical instability without a trigger, and reflects the high spatial variability of the snowpack at Big Sky. This supports observations by LaChapelle and Atwater (1961) that suggest the snowpack is near critical instability and requires just a trigger to initiate a deep slab.

Comey and McCollister (2008) also note explosives as the most common trigger for slab avalanches with crown depths deeper than 1.18m (4 feet), in a database from Jackson Hole ski area and the Bridger-Teton National Forest Avalanche Center in Wyoming. Explosives are often regarded as the ultimate stability test for deep slab avalanche hazard, and snowpits are seldom used to assess deep slab hazard at Big Sky, Montana due to the vast spatial variability that results from the interaction of the climate and complex terrain (Savage, 2006).

Large avalanches on one avalanche path in Switzerland were analyzed, primarily to determine the critical snow loads associated with these events and the ratio of the return period of the critical snow load to the return period of large avalanches (Schweizer et al., 2009). Thirty-four large events were identified, twelve of which reached the road in the runout of the avalanche path. Meteorological variables preceding the events were

compared between the 22 events that did not hit the road and the 12 events that did, and between the 34 days with large avalanches and 1,541 days without large avalanches. The variables and trends having the strongest association with the 34 large avalanches were: substantial snow loading, snow depth deeper than the terrain roughness, a slight increase in air temperature, and a snowpack with a weak base (Schweizer et al., 2009). A weak base existed for almost all avalanches that hit the road, but was also present for many of the other large avalanches, so it had weak discriminating power ($p=0.111$, U-test) between avalanches that reached the road and those that didn't. Alternatively, a change in snow depth over three days preceding the event had strong discriminating power ($p=.0008$, U-test) between large slides that did and didn't hit the road (Schweizer et al., 2009). The return period of the critical snow depth to initiate a large avalanche was found to be 2-5 times more frequent than the return period of a large avalanche in six additional paths, which leads to high (78-97%) false alarm ratios when using only critical snow depth to forecast large events (Schweizer et al., 2009).

Tracz (2012) used a record of 45,718 days with avalanches in western Canada to compare the characteristics of avalanches that failed on a persistent weak layer that was buried for more than three weeks, and had a crown depth greater than the $80^{th}$ percentile of crown depths in the region, with other avalanche types across different regions (Tracz, 2012). A different database, which did not contain information on weak layer age, was used to examine weather conditions that preceded naturally triggered (i.e., snowfall, wind, or temperature) avalanches with crown depths greater than 0.8m (Tracz, 2012). Only avalanches with a natural trigger were used in an attempt to strengthen the association with particular preceding weather patterns (Tracz, 2012). Precipitation

amounts were the best discriminator between days with these avalanches and days without these avalanches. However, the time period over which these amounts were significant varied regionally (Tracz, 2012). Preceding changes in air temperature and number of hours the air temperature was above freezing were also found to have a strong association with natural avalanches deeper than 0.8m in some regions. Tracz (2012) emphasizes the importance of rate dependent antecedent variables as many previous studies have suggested rapid changes to have a destabilizing effect on the snowpack (e.g., Jamieson et al., 2001).

Conlan et al. (2014) made snowpack observations and summarized preceding weather conditions at 41 "hard-to-forecast" avalanches in Western Canada. These were defined as avalanches that failed on a persistent weak layer, at some point (typically greater than one week) after the initial cycle of avalanches on that weak layer (Conlan et al., 2014). They found "hard-to-forecast" avalanches failing on depth hoar and facets to occur later after their burial date than those failing on surface hoar. Precipitation and positive temperature changes were shown to have a strong correlation with "hard-to-forecast" avalanches. However, precipitation amounts prior to "hard-to-forecast" avalanches were not much more than amounts accompanying storms that did not produce these avalanches. Warming prior to "hard-to-forecast" persistent avalanches was suggested to be associated with most snowstorms in the region (Conlan et al., 2014). Despite strong correlations between these avalanches and preceding precipitation and temperature change (Tracz, 2012; Conlan et al., 2014), a high false alarm ratio makes prediction of these avalanches difficult when using only these variables (Tracz, 2012). A more thorough multivariate approach is suggested to forecast deep slab avalanches

(Conlan et al., 2014). Separating early and late season events (Tracz, 2012), weak layer types, types of trigger, and regions are suggested to improve results for forecasting deep slab avalanches (Conlan et al., 2014).

This thesis focuses on deep slabs that failed on deep persistent weak layers, late in the season (February, March, or April), in one mountain range (Bridger Mountains, southwest Montana) over the course of recorded history (44 seasons). In addition to these more refined study characteristics, meteorological conditions during the time of weak layer formation (i.e., November, December, January) are analyzed in conjunction with weather directly preceding (1-7 days) deep slab avalanches.

## 2.4 Applied Avalanche Forecasting

Avalanche forecasting involves the prediction of the physical state of the snow cover with respect to stability. Snow cover instability, with regard to a specific triggering level, is what is actually forecast because instability leads to avalanches (McClung, 2002a). Applied avalanche forecasting (avalanche forecasting) requires the evaluation and prediction of snow cover instability followed by decisions that affect lives and property (McClung, 2002b). Therefore, avalanche forecasting is an analysis of risk, where risk is the probability of death or property loss. The goal is to use available information to obtain a perception of the distribution of instability throughout the snow cover that is as close to reality as possible (McClung, 2002a).

An avalanche forecast does not result instantly from the input of variables into a fixed formula. It is the result of continuous evaluation of data integrated over time. There are typically not explicit rules or steps to follow to forecast avalanches, but it is an

inherently inductive process that follows the scientific method (LaChapelle, 1980). Data are collected and hypotheses are formed about the state of the snow cover, and then they are confirmed and revised based on continuing conditions until a representation of snow cover stability is achieved with low uncertainty (LaChapelle, 1980). Information used to test hypotheses about instability is diverse. Avalanche forecasts consider data on weather, snow structure, avalanche occurrence, trigger type, climate and weather history, among many other data. All information becomes a part of the forecast regardless of its apparent significance (LaChapelle, 1980). Information that implies instability with low uncertainty (low entropy) is most valuable for achieving the goal of avalanche forecasting (LaChapelle, 1980; McClung, 2002a).

LaChapelle (1980, 1985) defines three categories of data: 1) meteorological data measured near the snow surface (i.e., snowfall amounts, wind speed and direction, temperature), 2) snow-cover structure data that indicate instability, and 3) snow mechanical data that reveal instability due to an applied load (McClung, 2002b). These categories are listed in decreasing order of informational entropy or potential information. Informational entropy essentially represents a relative level of uncertainty. It can also be viewed as potential information, or the amount of possible outcomes that didn't happen (LaChapelle, 1980). For example, meteorological data have many variables that could be interpreted to imply a variety of outcomes, while snow mechanics data (as viewed in this context) mostly refers to direct observations of snow stability or instability (e.g., avalanche occurrence, fracture propagation, and weak layer-slab structure), many of which represent binary outcomes of either stable or unstable (LaChapelle, 1980; McClung 2002b).

Information with high entropy can be transformed to minimize entropy, and essentially focus the data on the question: is the snow stable or unstable? Observing that the current temperature is 3°C or that the wind direction is northeast is not as easily interpreted as whether the temperature is above or below freezing, or whether the wind direction is or isn't increasing loading in starting zones (LaChapelle, 1980). Both high and low entropy information are useful in different forecasting settings and at different scales. An office-based forecast for a large region will rely more on high entropy data, and cope with the uncertainty through more data processing and iterative analyses. A field worker concerned with a smaller forecast area may rely more on low entropy information, which provides information on instability with less uncertainty (LaChapelle, 1980).

Iterative processes like season long re-evaluation minimize uncertainty (LaChapelle, 1980). LaChapelle (1980) states that redundancy in information minimizes uncertainty, and McClung (2002a) elaborates that redundant information may improve confidence in a forecast, but won't necessarily improve accuracy. However, an average of several independent assessments, or pieces of information, may improve the accuracy of a forecast (LaChapelle, 1980; McClung, 2002a).

McClung (2002b) describes two types of information that are necessary to forecast avalanches. The above three categories of data, which have informational entropy, are one type of information, referred to as singular information. Singular information, or data, is specific to the present and future situation. The other type of information is distributional data, which represent the general outcomes of situations in the past that are similar to the current situation. Distributional data represent a

forecaster's experience with terrain and avalanches. It may include integrated rules of thumb that are acquired through experience, and statistical models created with computers that compare similar situations from the past (McClung, 2002b).

Singular data are essentially required for an avalanche forecast, while distributional data may be extremely limited (non-existent) for persons with minimal (no) experience (McClung, 2002b). Avalanche forecasting has been described using Bayes Theorem (LaChapelle, 1980; McClung, 2002b), where the Bayes Rule is:

Posterior Probability α Prior Probability x Likelihood,

where singular data make up part of the Likelihood and distributional data are the Prior Probability. The Posterior Probability is the situation in question and later becomes a part of the Prior Probability (McClung, 2002b). A forecaster must consider the current state of the snow instability before making decisions about activity in avalanche terrain. Without singular data only the Prior Probability is available, which can lead to error and inaccuracy. Furthermore, a lack of distributional data can lead to bias. It is essential to objectively combine both types of data (singular and distributional) for avalanche forecasting (McClung, 2002b).

Human factors and perceptions introduce further uncertainty into avalanche forecasting (McClung 2002a). Personality traits are a function of life experiences and affect one's risk propensity, which defines the limits of acceptable risk for an individual. A decision is successful if the outcome falls within the boundaries of acceptable risk. An inaccurate avalanche forecast can lead to two types of error. Type I error occurs when an outcome falls above the acceptable level of risk. This error implies that an accident occurred because the snow was unstable when it was believed to be stable. Type II error

occurs when an outcome is below the lower limit of risk. This typically implies an excessively conservative decision was made that led to a lack of action (i.e., not opening ski terrain or transportation routes, or not ending avalanche warnings) after snow became stable. Individual and operational goals define the limits of acceptable risk. Acceptable limits of risk are determined by assessing the cost of either type of error. The cost for exceeding the upper limit (type I error) is injury, death, or property loss. Exceeding the lower limit (type II error) may result in lost revenue from closed transportation corridors or ski terrain, loss of credibility due to overly conservative warnings, or missed opportunities and feelings of underachievement (McClung, 2002a).

Human perceptions are improved and positively affected by targeted education and experience, and they are negatively affected by human bias that asserts prejudice and affects objective reasoning. The influence of bias can be dependent on the level of instability. In situations with obvious instability (i.e., widespread avalanche activity) bias has little effect on the clear decision to avoid avalanche terrain, since observation and reality closely match. It is in situations with moderate or considerable stability that bias can overshadow objective assessments of information, and bridge perception with reality through prejudice interpolation (McClung, 2002a). These situations are accompanied by much greater relative uncertainty.

The goal of an avalanche forecast is to predict the state of the snow cover, in regards to instability, with as little uncertainty as possible. Avalanche forecasting is essentially an analysis of risk, as it requires assessment of the snow cover instability in regards to the potential loss of life or property. The acceptable level of risk for an avalanche forecast varies depending on personal or operational goals. These desires will

add bias to an assessment of instability and can greatly contribute to the level of uncertainty in both positive and negative ways. An avalanche forecast is achieved by integrating weather and snowpack observations (singular data) with training and experience (distributional data). Higher entropy singular data typically requires more experience and some transformation, or processing, to interpret. However, high entropy data can be powerful and reduce uncertainty in a forecast when it is properly used. This thesis uses statistical tools combined with snow and meteorological observations to model the relationships between deep slab avalanches and summaries of meteorological observations prior to these avalanches. The results will help forecasters reduce uncertainty by suggesting useful summaries of raw meteorological data as a piece of an avalanche forecast.

## 2.5 Statistical Avalanche Forecasting

Statistical models relating meteorological data to avalanche activity have been developed to aid avalanche forecasters for over 40 years (Judson and Erickson, 1973; Hendrikx et al., 2014). A variety of statistical methods have been used (Table 3), which include univariate analysis (Perla, 1970; Jones and Jamieson, 2001), discriminant analysis (Bois et al., 1975; Bovis, 1977), linear regression (Floyer and McClung, 2003), binomial logistic regression (Romig, 2004), principal components analysis (Föhn et al., 1977), cluster analysis (Föhn et al., 1977; Obled and Good, 1980; McClung and Tweedy, 1994), nearest neighbors (Buser, 1983; Schweizer and Föhn, 1996; McCollister et al., 2003), classification trees (Davis et al., 1999; Hendrikx et al., 2005), and random forests (Guy and Birkeland, 2013). Models that result from these typically provide 70-85%

correct prediction of avalanche activity in the area where the data used to fit the model was derived (Table 3). This thesis uses univariate analysis, classification trees, and random forests to find meteorological factors associated with deep slab avalanches on persistent weak layers. Previous applications of these methods to avalanche forecasting are reviewed in detail in this section and are summarized in Table 3. Many of these applications focus on the explanation of factors that contribute to avalanches, while some are used to forecast avalanches using the input of new data.

### 2.5.1 Univariate and Multivariate Analysis

Univariate techniques were used on 20 years of avalanche data at Alta, Utah to test variables that were related to Atwater's (1954) ten contributory factors (Table 1) to avalanche hazard (Perla, 1970). Avalanche hazard, derived from the sum of all avalanches size 3, 4, or 5 that ran during or directly following a storm, was plotted against contributing factors for all storms on record. Each scatterplot was divided into four or five equal intervals on the x-axis. A probability of a hazard three or greater and probability of a hazard ten or greater were plotted in each interval. An increase in probability with an increase in a given factor was considered a first order contributor, and a substantial oscillation in probability with increase in a given factor was considered a second order contributor. Avalanche hazard was found to vary substantially with precipitation and wind direction, and marginally with temperature change. Maximum precipitation intensity was strongly related to avalanche hazard, supporting the importance of rate dependent variables on avalanche initiation (Perla, 1970). Wind speed and snow settlement had little effect on avalanche hazard (Perla, 1970).

Judson and Erickson (1973) used a similar approach with data from 23 avalanche paths in the Front Range of Colorado. Their univariate scatterplot analysis provided a storm index that correlated well with the number of avalanches (r=0.86) on the 23 paths. The storm index was derived from the sum of maximum precipitation intensity multiplied by a constant for excessive wind speed (Judson and Erickson, 1973). Other results of the univariate analysis were used to isolate 13 factors to be used in a multivariate discriminant function analysis on each of 10 paths. A discriminant function analysis finds the linear combination of variables that best separates two pre-defined groups. In this case, it was used to find the linear combination of meteorological variables that best discriminate between periods with stable snow and periods with un-stable snow, which were based on whether or not a path was active with control efforts during a storm (Judson and Erickson, 1973). Storms were considered to begin after 12 hours of consecutive precipitation and were over after 18-hours without precipitation. Three variables that were derived from maximum precipitation intensity, wind speed and direction, and temperature were used to develop discriminant functions for each path. The results of testing these functions in a forecasting mode (new data) were not published. However, the discriminant functions correctly classified 70-80% of the training data (Judson and Erickson, 1973).

Table 3 Characteristics and findings from previous studies that have applied statistical methods for avalanche forecasting and research.

| Reference | Location | Climate | Avalanche Type or Hazard Definition (i.e. Response variable) | Sample extent | Sample size | Method | Non-av day selection | Significant Predictors | Performance Assessment |
|---|---|---|---|---|---|---|---|---|---|
| Perla (1970) | Alta, Utah | Intermountain/ Coastal | Sum of sizes 3,4, and 5 during or directly following a storm | S. facing slopes above Alta village and road (similar paths) | 107 storms, 87 ram profiles (20 years) | Graphical univariate analysis of hazard probability against each variable | NA | Precip. intensity, wind direction, temperature change | NA |
| Judson and Erickson (1973) | Front Range, Colorado | Continental | Path active with control efforts following a storm? (Y/N) | 23 paths (mountain range) | 42-81 storms (8 years) | Graphical univariate analysis (Perla, 1970) (on all paths,) and discriminant function analysis (10 paths) | Used only days during or following a storm | Max. precip. intensity, wind speed and direction, temperature | 70-80% correct classification on training data |
| Bois et al. (1975) | Davos, Switzerland | Transitional | Days with dry avalanches, wet avalanches, or no avalanches (natural only) | (Mountain Region) | 10 years (1961-1970) | Discriminant analysis | Random sample equal to the number of avalanche days in each month | New snow, temperature, and wind | NA |
| Bovis (1977) | San Juan Mtns, Colorado | Continental/ Intermountain | Days with and without avalanches (stratified by avalanche size, count, and type) | 150 paths (mountain range) | (2 years) | Discriminant analysis (to separate non-avalanche days from different strata of avalanche days) | Random sample equal to the number of avalanche days | Max 6 hour precip intensity over 24 hours, total SWE over 3 days, wind redistribution after high precip (dr avys). Avg. and max. 2,3, and 5 day 2 hour temperatures (wet avys). | 0-57% misclassifciation |
| Föhn et al. (1977) | Davos, Switzerland | Transitional | Days with avalanches and days without avalanches | (Mountain Region) | (~20 years) | Conventional, principal components, cluster analysis | NA | New snow, wind speed, air temp., radiation, depth of active slab | 70-80% of days classified well |
| Obled and Good (1980) | Davos, Switzerland | Transitional | Days without avalanches and days with natural avalanches | ~100km^2 (mountain region) | 1683 days, 307 avalanche days (14 years) | Discriminant analysis, cluster analysis, nearest neighbors | Every fifth day (without avalanches) selected | Most | 3-27% Misclassifciation |
| Buser (1983) | Davos, Switzerland | Transitional | All avalanches | (Mountain Region) | (20 years) | Nearest neighbors | NA | All | Informational tool |
| Buser (1989) | Davos, Switzerland | Transitional | All avalanches | (Mountain Region) | Fit from 20 years (tested on 2 years) | Nearest neighbors (Buser, 1983) | NA | All | Informational tool (Qualified as useful) |
| McClung and Tweedy (1994) | Kootenay Pass, British Columbia | NA | Dry avalanches, moist-wet avalanches, and days with each type and AAI > 3 or >10 | (Road corridor/ Mountain region) | 10 years (1981-1990) | Discriminant analysis, Bayesion statistics, cluster analysis, and nearest neighbors | Randomly selected equal to the number of avalanche days | All | 70-88% correct classification across groups |
| Schweizer and Föhn (1996) | Davos, Switzerland | Transitional | Avalanche hazard (elevation and aspect) | ~50 km^2 (forecast area) | 1361 days (9 years) | Expert systems and nearest neighbors | All included | Sum of new snow, penetration, snow depth, Snow drift, snow stability index, depth to weak layer | 42-75% Performance |

Table 3 (Continued).

| Reference | Location | Climate | Avalanche Type or Hazard Definition (i.e. Response variable) | Sample extent | Sample size | Method | Non-av day selection | Significant Predictors | Performance Assessment |
|---|---|---|---|---|---|---|---|---|---|
| Davis et al. (1999) | Alta, Utah | Intermountain/ Coastal | Avalanche day (y/n), maximum avalanche size, or sum of avalanche sizes on a given day | Ski area | 1895 days, 669 avy days (1983-1994) | Classification and regression trees | All included | Lagged snowfall and SWE, wind drift parameter, starting snow depth of year, cumulative vapor pressure difference | 70-97% correct classification, R^2 of 0.84 for regression tree |
| Davis et al. (1999) | Mammoth, California | Coastal/ Intermountain | Avalanche day (y/n), maximum avalanche size, or sum of avalanche sizes on a given day | Ski area | 890 days, 160 avy days (1989-1998) | Classification and regression trees | All included | Lagged snowfall and SWE, wind drift parameter, starting snow depth of year, cumulative vapor pressure difference | 14-98% correct classification, R^2 of 0.88 for regression tree |
| Jones and Jamieson (2001) | British Columbia | Maritime-coastal transitional | Daily maximum size of skier triggered dry slab avalanches | ~5000km^2 (Heli-skiing operation) | 1292 days (1990-2000) | Univariate (boxplots, spearman rank correlations), classification trees (cross-validated) | All included | Previous avalanche activity, new snow or precipitation, height of snowpack, days since December 1st, relative humidity, barometric pressure | 57% correct classification on data used to fit the tree, 37% on data withheld from fitting (1998-2000). 81% on a tree fit from all data (1990-2000) |
| Floyer and McClung (2003) | Bear Pass, British Columbia | Maritime | Avalanche day (y/n) | Paths along road (length of road unknown) | 17 years (1985-2002) | ANOVA, Canonical discriminant analysis and discriminant functions | equal number of non-avy days as avy days | New precipitation, current temperature, snow depth, foot penetration, current temperature trend | 72% correct classificaion on dataset not used to fit model, 75% on data used to fit model. 70-79% for smaller areas grouped by similar characteristics |
| McCollister (2003) | Jackson Hole, Wyoming | Intermountain/ Continental | Probability of avalanches for a given area and given weather pattern | (Ski area) | 23 years (1978-2001) | Nearest neighbors | All included | More new snowfall increased avalanche probability at all scales. Wind speed increased probability in individual paths | Informational tool |
| Schweizer and Jamieson (2003) | Canada and Switzerland | Various | Stable vs unstable snowpack profiles | NA | 424 snow profiles | Univariate (Mann-Whitney U-test), disriminate analysis, classification trees (cross-validated) | 204 Stable, 220 Unstable | Rutschblock score, grain size difference, slab hardness, weak layer grain size and hardness | 68% correct classification for the discriminate analysis, 78% for the classification tree |
| Hendrikx et al. (2005) | Milford Road, New Zealand | Maritime | Avalanche day (size $\geq$ 2.5 or sum of sizes $\geq$ 10), or non-avalanche day | ~40km road | 148 avy days, 10 years (1993-2002) | Classification trees (cross validated) | chosen equal to the number of avalanche days in each year | 72 hr wind drift parameter, avg. 72-hr wind speed, avg. 12-hr air pressure, max. 72-hr air temperature, and avg. 72-hr snow depth | 85% overall correct classification and 79% POD. 78% overall and 86% POD with unequal misclassification costs |
| Schweizer et al. (2009) | Davos, Switzerland | Transitional | Large avalanches on one path | (One path) | 34 avalanches (58 years) | Mann-Whitney U-test, Classification trees | Days with precipitation $\geq$ 10cm (1,541 days) | Cumulative new snow and snow depth | POD: 67-100%, FAR: 78-97%, RPC: 77-89%, TSS: 29-54% |
| Baggi and Schwiezer (2009) | Dischma, Switzerland | Transitional | wet slab avalanche days vs. non-avalanche days | (Study area) | 27 wet slab days, 31 non-avalanche days (20 years) | Mann-Whitney U-test, Classification trees (cross-validated) | Chosen to equal the number of wet slab days in any month | capillary barriers, days since snowpack went isothermal, 3 day sum of positive air temperature | 88% for not cross validated tree, 71% for cross validated tree |
| Peitzsch et al. (2012) | Glacier National Park, USA | Intermountain | wet slab and glide avalanche days | 80km road | 88 days with $\geq$ 1 wet slab or glide avalanche (8 years) | Wilcoxin rank sum test, Classification trees (cross-validated) | Equal to the number of wet slab or glide days in any month | maximum daily air temperature, mean daily air temperature, and the change in snow depth over five days | 73% correct classification, POD: 69%. Correct classification of 10 of 11 wet slab days not used to fit the tree |

Table 3 (Continued)

| Reference | Location | Climate | Avalanche Type or Hazard Definition (i.e. Response variable) | Sample extent | Sample size | Method | Non-av day selection | Significant Predictors | Performance Assessment |
|---|---|---|---|---|---|---|---|---|---|
| Hendrikx et al. (2014) | Seward Highwy, Alaska | Maritime | Avalanche days (≥ 1 avalanche over 90% of its path length) vs. non-avalanche days | 200km road | 559 avalanche days, 394 non-avalanche days (28 years) | Cross validated classification trees | Equal to number of avalanche days in each month and year | 72-hour sum of water, 24-hour high temperature, 72-hour high temperature | RPC: 72%, POD: 77%, FAR: 24%, TSS: 47% on data used to fit tree. RPC: 62-72%, POD: 46-61%, FAR: 62-67%, TSS: 16-44% when used to predict data not used to fit tree |
| Marienthal (2014) | Bridger Bowl, Montana | Intermountain/ Continental | Days and seasons with deep slabs (>0.9m) on persistent weak layers after Feb. 1 | Ski area | 76 days (24 avy seasons, 20 non-avy seasons) | Cross validated classification trees | Equal to number of avalanche days in seasons w/ avys | Dry early season with snow depth conducive to weak layer development, high cumulative precip. prior to days with deep slabs, sustained warming prior to deep wet slabs | Dry early season with snow depth conducive to weak layer development, high cumulative precip. prior to days with deep slabs, sustained warming prior to deep wet slabs |

Bois et al. (1975) mention that the large number of days without avalanches in most datasets violates assumptions of normality in many traditional statistical analyses. For example, response variables in regression analyses are often a count or sum of avalanche sizes, so using all days in a season makes for a relatively large number of days with a zero value. Judson and Erickson (1973) reduced the number of non-avalanche days by only considering days during storms. A variety of other approaches have been used to address the statistical problem of having too many days without avalanches (Table 3).

Bois et al. (1975) used meteorological data and avalanche records from Davos, Switzerland in a discriminant analysis to determine meteorological conditions that best distinguish between days with dry avalanches, days with wet avalanches, and days without any avalanches. Days without avalanches were randomly selected so that there was an equal number of days without avalanches and days with avalanches in each month. Only natural avalanches were considered because artificial control was not systematically implemented. New snow, temperature, and wind variables were the best variables to separate dry avalanche and non-avalanche days. Temperature, radiation, and previous avalanche activity were found to be the variables that best separated wet avalanche days from non-avalanche days (Bois et al., 1975).

Discriminant analysis was used to find meteorological and snowpack variables that distinguish between avalanche and non-avalanche days on 150 paths in the San Juan Mountains of Colorado (Bovis, 1977). Days with dry and wet avalanches were treated in separate discriminant analyses due to the difference in factors that contribute to their formation. A random sample of days without avalanches was chosen equal to the number of days with avalanches. Variables were integrated from two, three, and five days prior to

an event day. Dry snow avalanche days were best separated from non-avalanche days by six-hour maximum precipitation intensity over 24 hours preceding the event day. Integrating meteorological variables from three days prior to an event day was found to best discriminate between avalanche and non-avalanche days. Total snow water equivalent over a three-day period was secondary to the six-hour maximum precipitation intensity (Bovis, 1977). Avalanche days were stratified by the magnitude and number of events. There was high misclassification (18-53%) when all avalanche days were used, which generally improved with stratification. This suggests that factors that contribute to days with one or a few small slides differ from those that contribute to days that have an avalanche "cycle" (Bovis, 1977). The best separation was found between non-avalanche days and days with avalanches greater than or equal to size three. This separation was defined by wind redistribution when there were high precipitation totals over preceding days. Wet avalanches were better separated from non-avalanche days than dry avalanches were. Average and maximum two-hour air temperature over two, three, or five days was the best discriminator for wet avalanche days, followed by three and five day precipitation totals. Stratifying wet slabs by quantity and magnitude reduced misclassifications. Ultimately, prediction was found to improve with stratification of avalanche days by type and magnitude (Bovis, 1977).

Statistical methods were evaluated alongside conventional forecasting methods (subjective analysis of avalanche and weather variables) using a record of avalanches and meteorological data from Davos, Switzerland (Föhn et al., 1977). Principal components analysis and a stepwise variable selection process based on Wilk's criterion were used with popular meteorological variables (similar to those used in previously described

studies) in one method, and were used with regional-scale meteorological variables (e.g., 700mb height) in another method. A dynamic cluster analysis was performed as well, which attempts to stratify avalanche days into groups based on within group similarities of meteorological conditions. They found the conventional method, the principal components analysis using common meteorological variables, and the cluster analysis to perform similarly well, with 70-80% of days classified well. The other statistical methods performed relatively poor (Föhn et al., 1977).

Obled and Good (1980) evaluated three statistical methods for forecasting avalanches using meteorological and avalanche records from Davos, Switzerland. They used only days with natural avalanche activity as avalanche days because there was no systematic avalanche control performed in the region. Data was stratified into two samples defined by the periods January to February and March to April. Stratification by avalanche type, number, and magnitude (e.g., Bovis, 1977) was not considered due to uncertainties that result from under-observation (Obled and Good, 1980). The first method was a discriminant analysis between avalanche and non-avalanche days in each time period. They chose every fifth day that did not have avalanche activity to comprise the population of days without avalanches. They found more new snow and lower air temperatures were associated with avalanche days in January and February. In March and April warm sunny days were associated with avalanche activity while warm snowy days were associated with no avalanche activity. The latter suggests that using only avalanche and non-avalanche day groups is insufficient even though previous studies have found similar conclusions using more stratification (Obled and Good, 1980).

The second analysis performed by Obled and Good (1980) was a cluster analysis that separated avalanche days into three or four groups of different combinations of meteorological conditions. Non-avalanche days were then assigned to one of the groups, which results in similar days that are either avalanche or non-avalanche days. A discriminant function was applied to each group in an attempt to separate avalanche and non-avalanche days within each group. A nearest neighbors approach was used in the third statistical method. More detail on this is given in section 2.5.2, below. The first method, using discriminant analysis, was considered to be robust and have similar performance to the other two methods. The cluster analysis and nearest neighbors methods provided additional interpretive information to forecasters, but could be improved through larger sample sizes (Obled and Good, 1980).

Floyer and McClung (2003) used ANOVA and canonical discriminant analysis (CDA) to find significant weather and snowpack variables that were later used in discriminant functions for forecasting avalanche days on Bear Pass in British Columbia, Canada. They defined four areas of similar terrain based on aspect, size, and elevation, and then compared discriminant functions for these individual areas to functions for the whole pass in order to test if discrimination between non-avalanche days and avalanche days was better for homogenous areas. A number of non-avalanche days equal to the number of avalanche days were selected randomly, and a random sample of 35% of all days was withheld from the dataset for model testing (Floyer and McClung, 2003). Variables that were found to be significant and used for forecasting avalanche days were new precipitation, current temperature, snow depth, foot penetration, and current temperature trend. The discriminant function created from these variables had an overall

classification rate of 72% on the dataset that was initially withheld from the analysis. This was slightly below the classification rate (75%) on the dataset that was used to create the discriminant function (Floyer and McClung, 2003). Discriminant functions for the four individual areas had similar variables as the function for the entire area, with a few differences. Current temperature trend was not used, minimum temperature replaced current temperature for one area, and two areas used new snow board depth. Classification rates for these functions were higher than the rates for the entire area in three of the four individual areas. Overall classification rates for the four individual areas ranged from 70-79% (Floyer and McClung, 2003).

### 2.5.2 Nearest Neighbors Analysis

Nearest neighbors analysis finds a specified number of days that are the most similar to a given day and calculates an avalanche probability from the proportion of similar days that had avalanches.

A nearest neighbors approach was used on the same data as Obled and Good (1980) by comparing a given avalanche day to the ten most meteorologically similar days, and deriving a probability from the number of those days that had avalanches (Buser, 1983). Meteorological variables were weighted differently if the user felt the resulting similar days were different than the actual situation. For avalanche days tested with the nearest neighbors approach by Buser (1983), avalanches on similar days (found using nearest neighbors) were on similar aspects and elevations and from similar triggers as the avalanches on the day being tested. The largest limitation of this approach was when non-avalanche days could not be distinguished from days when avalanches

occurred without being observed or occurred one or two days prior to the day that they were recorded (Buser, 1983). This approach allows users to examine similar days individually, which allows them to directly compare them with the actual situation. This is beneficial because forecasters are often more familiar with actual situations than with probabilities (Buser, 1983). Furthermore, nearest neighbors models allow those with sufficient understanding of avalanches, but minimal local experience, to quickly gain knowledge of local patterns (McCollister et al., 2003).

Avalanche control workers used the nearest neighbor model (Buser, 1983) operationally over two seasons in the area that it was developed (Buser, 1989). Workers still relied on their experience and expertise, but found the model useful to provide more detail about a familiar situation. Although the proportion of nearest neighbor days having avalanches could often be low (e.g., 10%-40%), the detailed record allowed workers to see how days with avalanches could be more or less similar to today's situation (Buser, 1989). The model was regarded to give information or prompt workers to examine previously overlooked areas. In examples of specific avalanche days, nearest neighbor days often contained days with avalanches on similar aspects and from similar triggers (Buser, 1989).

McCollister et al. (2003) used nearest neighbors at Jackson Hole Mountain Resort in Wyoming to create avalanche probabilities for a range of new snowfall, wind direction, and wind speed. Probability distributions were derived for the entire ski area, individual avalanche paths, groups of nearby paths, and paths that were grouped by aspect and elevation. Higher amounts of new snowfall increased avalanche probabilities for the ski area, all groups, and individual paths. However, probabilities increased with

wind speed in some paths while it decreased in other paths. Probabilities increased slightly with wind speed over the entire ski area and groups of paths. This displays the variability of the effect of wind speed at the individual path scale, while these effects are essentially averaged out over larger areas (McCollister et al., 2003).

Avalanche probabilities were derived from discriminant functions, Bayesian statistics, cluster analysis, and nearest neighbors on Kootenay Pass in British Columbia (McClung and Tweedy, 1994). They used these techniques to separate days without avalanches, days with dry avalanches, days with moist-wet avalanches, days with dry avalanches with an avalanche activity index (AAI ) $\geq$ 3 or $\geq$10, and days with moist-wet avalanches with an AAI $\geq$ 3 or $\geq$10. They had 70-88% correct classifications across all groups (McClung and Tweedy, 1994).

Schweizer and Föhn (1996) used an expert system approach to improve the performance of a nearest neighbor model. An expert system essentially simulates the expert's decision-making process by using stepwise rules defined by experts. They developed a nearest neighbor model (DAVOS) that used 13 derived and measured weather and snow-cover parameters to provide a degree of hazard and location (i.e., elevation and aspect). They also developed the MODUL model, which gives the degree of hazard from a series of rule-based decisions (Schweizer and Föhn, 1996). Model performance was based on the model correctly classifying the degree of hazard as defined by an expert's verified degree of hazard (i.e., degree of hazard determined in hindsight). Different versions of the DAVOS model achieved an average performance of 60% (42-75%), and the MODUL model achieved a performance of 73%. The addition of an expert system process, as well as more snow-cover parameters, improved the model

performance. Although snow-cover parameters are known to be an important factor in determining avalanche hazard, they had seldom been used in prior statistical forecasting models (Schweizer and Föhn, 1996). A slightly different nearest neighbors model (NEX-MOD by Meister, 1991), which is based off the model used in Buser (1989), had a performance of 38-47% over the same three years that DAVOS and MODUL were tested (Schweizer and Föhn, 1996).

### 2.5.3 Classification Trees

Classification trees have been applied more recently for statistical avalanche forecasting (e.g., Davis et al., 1999; Hendrikx et al., 2014) because they are essentially non-parametric and largely insensitive to underlying distributions of data (Breiman et al., 1993). Furthermore, they are useful for both prediction and explanation, and they are often more clear and understandable than other statistical methods (Davis et al., 1999; Hendrikx et al., 2005). Classification trees split a dataset into increasingly homogenous groups of observations, in this case either events or non-events, based on a threshold value of explanatory variables. Homogenous groups have all observations of one type. Non-homogenous groups can be further divided using threshold values of other explanatory variables (Breiman et al., 1993).

Davis et al. (1999) used classification trees and regression trees to explore associations between 31 conventional and derived meteorological and snow-cover variables and days that had dry avalanches with any type of trigger at Alta, Utah and Mammoth Mountain, California. They used classification trees to assess the importance of variables in distinguishing between avalanche days and non-avalanche days, or

between maximum relative size (R-size) classes for a given day. Regression trees were used to assess the importance of variables in determining the sum of avalanche sizes on a given day (Davis et al., 1999). They used over fit trees in an attempt to explain the underlying structure of the data rather than for prediction, which would ideally be done using a pruned tree (Davis et al., 1999). They limited their over-fit trees to 100 or less splits. Probabilities of correct classification from over fit trees for a given response variable ranged from 0.14 (maximum R-size 5 at Mammoth) to 0.98 (avalanche day at Mammoth). The probability of correct classification of an avalanche day at Alta was 0.97, and the probability of correct classification of maximum size ranged from 0.14 to 0.95 at Mammoth and 0.70 to 0.94 at Alta. Regression trees for the sum of sizes had an $R^2$ of 0.84 for Alta and 0.88 for Mammoth. Probabilities of correct classification would be lower in pruned trees being used for prediction (Davis et al., 1999).

The importance of explanatory variables was determined from their ability to reduce misclassification accuracy when used as primary or surrogate splitting decisions (Davis et al., 1999). Similar to previous studies, lagged variables of snowfall depth and SWE (i.e., 2 and 3 day total snowfall or SWE) ranked highest in importance. A derived wind drift parameter (i.e., the product of average wind speed to the fourth power and total snowfall) over 2 and 3 days showed the highest importance of any derived variable in reducing misclassifications, and had scores slightly below lagged SWE and snowfall when considering all conventional and derived variables. Furthermore, starting snow depth of the year was the most important of between year and between storm variables, followed by accumulated vapor pressure difference (Davis et al., 1999).

Jones and Jamieson (2001) assessed the relative importance of 16 measured and 14 derived meteorological variables, and two previous avalanche activity variables for forecasting the daily maximum size of skier triggered dry slab avalanches at a helicopter skiing operation in British Columbia. They used boxplots and spearman rank correlations to find variables that were strongly correlated with maximum avalanche size. A classification tree was fit that ceased splitting when there were five or less observations in a group. Days with data available from 1990-1998 were used to create the tree, and those from 1998-2000 were used to validate the tree. The response variable for the trees consisted of three categories to define a day. These were days without avalanches, days with maximum destructive size (D-size) less than 1.5, and days with maximum D-size greater than 1.5. The classification tree had an overall correct classification of 57% on the data used to fit the tree. The overall correct classification on the data withheld for testing was 37%, with only 26% of non-avalanche days correctly classified. A tree was fit to the entire dataset (including all years from 1990-2000), and had an overall correct classification of 81%, with 77% of non-avalanche days correctly predicted and 96% of avalanche days correctly classified. The tree with the entire dataset did not have any avalanche days predicted as non-avalanche days (Jones and Jamieson, 2001). Meteorological variables found to have the strongest association with skier-triggered dry slab avalanches included previous avalanche activity, new snow or precipitation, height of snowpack, days since December 1$^{st}$, relative humidity, and barometric pressure. Change in barometric pressure since the previous day was the highest ranked predictor in the classification tree, followed by height of snowpack (Jones and Jamieson, 2001).

Schweizer and Jamieson (2003) compared snowpack properties from stable snow profiles to properties of unstable snow profiles in Canada and Switzerland. Unstable profiles were collected adjacent to slopes that had a skier-triggered avalanche, and stable profiles were from slopes that had been skied without an avalanche being triggered. They used the Mann-Whitney U-test to compare the distributions of each individual variable in stable and unstable profiles. Discriminate analysis and classification trees were used to explore which snowpack properties could best classify or discriminate between stable and unstable profiles. Rutschblock score was found to best separate between stable and unstable profiles in both multivariate analyses (Schweizer and Jamieson, 2003). The discriminate analysis had an overall correct classification of 68% (based on the jackknifed procedure where one observation is removed and replaced one at a time). The discriminate analysis had a classification score of 64% when rutschblock score was removed. The classification tree (using a minimum splitting criteria of at least a 0.02 proportion reduction in error) correctly classified the data used to fit the tree 78% of the time, and 74% of the time without using the rutschblock score (Schweizer and Jamieson, 2003).

Hendrikx et al. (2005) used pruned classification trees to predict avalanche days versus non-avalanche days on the Milford Road in New Zealand. They defined avalanche days as days with a maximum avalanche size class $\geq 2.5$, or sum of size classes $\geq 10$. They generated an equal number of non-avalanche days by selecting an equal number of days without avalanches as days that had natural avalanches in each year. In addition to this set of days, they included all (47) days that had avalanches but did not meet the criteria for being an avalanche day, as non-avalanche days. This resulted in 148

avalanche days and 195 non-avalanche days. They used both wet and dry avalanches from both natural and artificial triggers between May 30th and December 20th. An over-fit classification tree was created using 48 meteorological variables to split data into increasingly homogenous groups of avalanche days or non-avalanche days (Hendrikx et al., 2005). They used a 10-fold cross validation technique (Breiman et al., 1993) to prune the over-fit tree to the optimal size for prediction.

The tree that resulted from the cross validation had five splits. The first split was made on a 72-hour wind drift parameter similar to the one used by Davis et al. (1999) (i.e., the product of total snowfall and average wind speed to the fourth power), and which also accounted for air temperature. The following splits (in order of importance) were made on average 72-hour wind speed, average 12-hour air pressure, maximum 72-hour air temperature, and average 72-hour snow depth (Hendrikx et al., 2005). This tree gave an overall classification accuracy of 85% with a probability of detection (POD) of 79%. The 21% of incorrectly predicted (false negative) avalanche days may be more significant for operational forecasting purposes, so a misclassification cost was introduced to prune to a tree that reduced false negatives. The misclassification cost essentially penalizes trees more for false negatives than for false positives. The resulting tree, with adjusted misclassification costs, contained only the first two splits of the previous tree. The overall accuracy of this tree was lower at 78%, but had a POD of 86%, therefore lowering false negatives to 14% at the expense of increasing false positives (Hendrikx et al., 2005). The trade of higher overall accuracy for a more conservative prediction may be beneficial to operational forecasting, while the larger tree (without

misclassification costs) provides more insight about the underlying processes (Hendrikx et al., 2005).

Schweizer et al. (2009) used classification trees to examine the predictive ability of seven snow variables, a temperature change variable, and a binary variable for existence of a weak base to forecast large avalanches on one path in Davos, Switzerland. They considered 34 large avalanches that occurred on the path over 58 years. Daily meteorological values were used from all days over the 58 years that had new snowfall greater than or equal to 10cm with the exception of days immediately following large avalanches (Schweizer et al., 2009). This resulted in 1,541 days without avalanches that met these criteria. Classification trees were used to find splits on individual, and combinations of, variables that were found to be significantly different between two categories based on a Mann-Whitney U-test (Schweizer et al., 2009). They optimized splits to separate three binary categories: 1) days with large avalanches and days without avalanches, 2) days with large avalanches hitting the road and days without avalanches, and 3) days with large avalanches hitting the road and days with large avalanches not hitting the road. Predictive performance was partially described using unweighted average accuracy (RPC) (the average of the probability of detecting events (POD) and the probability of detecting non-events (PON)), the POD, and the FAR (proportion of days classified with large avalanches that did not actually have avalanches). Trees used to discriminate between days with large avalanches and days without avalanches reached an RPC of 77%. This was using a tree with splits on new snow over the preceding five days and snow depth. The POD for this tree was 68% and the FAR was 90%. The trees used to split days with avalanches hitting the road and days without avalanches had a slightly

better forecasting ability, with higher RPCs (77%-89%). However, the POD ranged from 67%-100% with FAR increasing with POD (78%-97%). For example, the highest POD was 100% with an FAR of 97% and RPC of 88%, and the lowest FAR was 78% with a POD of 67% and RPC of 82% (Schweizer et al., 2009).

Baggi and Schweizer (2009) used a classification tree for exploratory analysis of meteorological and snowpack variables related to wet slab avalanche activity in Dischma, Switzerland. They used seven of 36 variables that were significantly different between wet slab avalanche days and non-avalanche days (based on Mann-Whitney U-test) as inputs for classification trees. There were 31 non-avalanche days and 27 wet slab avalanche days from 20 years of data that were used to train the trees. Non-avalanche days were randomly chosen with the goal of having an equal number of non-avalanche days as avalanche days in any month. A tree that was not cross-validated had an accuracy of 88%, and a 10-fold cross validated tree had an accuracy of 71%. The variables shown to be most prominent from the cross validation were days since the snowpack reached isothermal state, index of capillary barriers, and the three day sum of positive air temperatures (Baggi and Schweizer, 2009).

A classification tree was trained and used to forecast days with wet slab and glide avalanche activity along Going-to-the-Sun-Road (GTSR) in Glacier National Park, USA (Peitzsch et al., 2012). There were 88 days from 8 years of data (2003-2010) having at least one wet slab or glide avalanche that were used to train the classification tree, and 11 days in 2011 that were used to test the tree for prediction. Meteorological variables that were significantly different between avalanche and non-avalanche days in a univariate analysis (Wilcoxon rank sum test) were used for input into a classification tree. A similar

approach to Hendrikx et al. (2005) was used where an over-fit tree was created and then pruned to a statistically defendable size for prediction using 10-fold cross validation (Peitzsch et al., 2012). Cases with missing values for meteorological variables had missing values estimated with the mean of that variable in order for incomplete cases not to be removed from the classification tree. A number of non-avalanche days were chosen that was equal to the number of avalanche days in each month. The final tree used for prediction split avalanche days from non-avalanche days using maximum daily air temperature, mean daily air temperature, and then the change in snow depth over five days. The percent of correctly classified cases was 73% with a POD for avalanche days of 69%. When used for days outside the training period (i.e., the 2011 season), the tree correctly classified 10 of the 11 days that had avalanches and misclassified 13 non-avalanche days as avalanche days. The number of non-avalanche days in 2011 was reduced to equal the number of avalanche days in each month of the 2011 season. Only 3 of these were misclassified as avalanche days, resulting in 82% correctly predicted days and a POD of 91% for the 2011 season (Peitzsch et al., 2012).

Hendrikx et al. (2014) used cross-validated classification trees (e.g., Hendrikx et al., 2005; Peitzsch et al., 2012) to predict avalanche days on the Seward Highway in Alaska. Trees were trained from 28 years (1983-2011) of meteorological and avalanche data. They used 30 derived and accumulated (i.e., 24, 48, and 72 hour) temperature, snow, and wind parameters to train trees that distinguished between avalanche and non-avalanche days. Avalanche days were defined as days that had at least one avalanche run 90% or more of its path length, where 100% of the path length is the highway. Non-avalanche days were chosen so there were an equal number of avalanche days as non-

avalanche days in each month and year. Days that were chosen as non-avalanche days and had extensive activity, but did not meet the criteria for an avalanche day, were removed. There were 394 non-avalanche days and 559 avalanche days. A second tree was trained using higher misclassification costs for false negative than for false positives, similar to Hendrikx et al. (2005). Both trees' predictive ability were assessed using the initial training data and their "true predictive ability was assessed using them in hindcast mode (with observed meteorological data as input) for the 2011-12 season, and in hindcast and a 24-hour forecast mode (using observed data and forecasted 24-hour data as input) for the 2012-13 season (Hendrikx et al., 2014).

The tree with equal misclassification costs used the 72-hour sum of water, the 24-hour high temperature, and the 72-hour high temperature. It had an RPC of 72%, POD of 77%, and a FAR of 24%. The tree with unequal misclassification costs used the 72-hour sum of water and the 24-hour average high temperature. It had a RPC of 64%, POD of 94%, and a FAR of 33%. For the tests of "true" predictability using data from 2011-12 and 2012-13, the tree with equal costs had RPCs ranging from 62-72%, POD ranging from 46-61%, and FARs from 62-67%. The tree with unequal costs had RPCs that ranged from 58-61%, PODs from 73-100%, and FARs from 77-83% (Hendrikx et al., 2014). The model here did not improve with the use of unequal misclassifications costs, unlike in Hendrikx et al. (2005). Using forecasted 24-hour metrics instead of observed (hindcasted) metrics did not greatly affect the strength of the model. Overall, the model performed well in the 2012-13 season (POD of 61% and RPC of 67%), but performed more poorly in the 2011-12 season (POD of 46% and RPC of 62%). Using the trees to predict days not used in the training dataset resulted in much larger FARs than those obtained from the

training data. This is also seen in Schweizer et al. (2009), and it is partly attributable to the larger number of non-avalanche days that occur when used operationally, as opposed to the equal number of days in the training data (Hendrikx et al., 2014). Forecasters regard the trees as most useful for forecasting events that have a higher uncertainty of occurrence (i.e., avalanches that occur some days after the main (obvious) storm or cycle), and found the concurrence of both trees on a non-avalanche day to help identify the end of a cycle (Hendrikx et al., 2014).

### 2.5.4 Random Forests

Random forests have been regarded as a more robust method to assess explanatory variables' predictive power and importance (Breiman, 2001). The random forest method iteratively grows classification trees while successively withholding (with replacement) random variables and subsets of data. This method gives a more accurate measure of predictive model performance at the expense of interpretability.

Guy and Birkeland (2013) used random forests to explore the importance of specific terrain parameters for predicting the location of potential trigger locations in steep couloirs. They used an overall success rate and true skill statistic (TSS) to quantify model performance of their random forests. The TSS is a measure of the model's ability to differentiate between events and non-events, where a value of 1 indicates the model can correctly identify all outcomes, a value of zero indicates that the model cannot differentiate between events and non-events, and a value of -1 means that all outcomes are incorrectly classified. They had success rates from 62-80% and TSS from 0.17-0.38 when using random forests (Guy and Birkeland, 2013).

This thesis will use random forests, classification trees, and univariate analysis to examine the association between meteorological metrics and avalanches that fail on deeply buried persistent weak layers (see section 3.4.4). A univariate analysis (Wilcoxon rank-sum test) will be performed to see which individual meteorological metrics are different between days with and days without, and seasons with and seasons without, deep slabs on persistent weak layers. Random forests and both over-fit and cross-validated classification trees will be trained that use meteorological metrics to split days with from days without deep slabs on persistent weak layers, as well as to split seasons with and seasons without this type of avalanche.

# 3. METHODS

## 3.1 Study Area

Deep slab avalanches that failed on persistent weak layers late in the season (February 1st or later) were studied at Bridger Bowl ski area. Bridger Bowl is located in southwest Montana, 24 km northeast of the town of Bozeman in the Bridger Mountains (Fig. 5). The ski area is on the east side of a north-south oriented ridgeline. Sixty percent of the ski area is avalanche terrain as classified by slope angle (greater than 30°), with elevations between 2075m and 2677m, and aspect ranging from North to East to South with an average aspect of 98° (ESE). Avalanche starting zones at Bridger Bowl are primarily between 2400m and 2677m. Much of the higher angle terrain is steep, rocky, and unable to retain snow, so there are very few starting zones steeper than 45°.

The Bridger Mountains are made up of steeply dipping (60-80° NE) beds of sedimentary rocks that have been modified by Quaternary erosional processes (Skipp et al., 1999). Glaciers, landslides, and streams are responsible for shaping the steep cirques and valleys of Bridger Bowl. The general stratigraphy near Bridger Bowl is comprised of Mississippian limestones overlain by (bottom to top) Permian and Pennsylvanian sedimentary rocks, Jurassic sedimentary rocks, and Cretaceous sandstones and shales (McMannis, 1955). These layers of rock, which were once deposited horizontally, have been tilted to a nearly vertical orientation during various tectonic events (Skipp et al., 1999). The more resistant limestone layers make up the crest of the ridge and the steep, high elevation, terrain inside Bridger Bowl, while the more erosive sandstone and shale comprise the lower angle foothills on the east side of the range (Skipp et al., 1999).

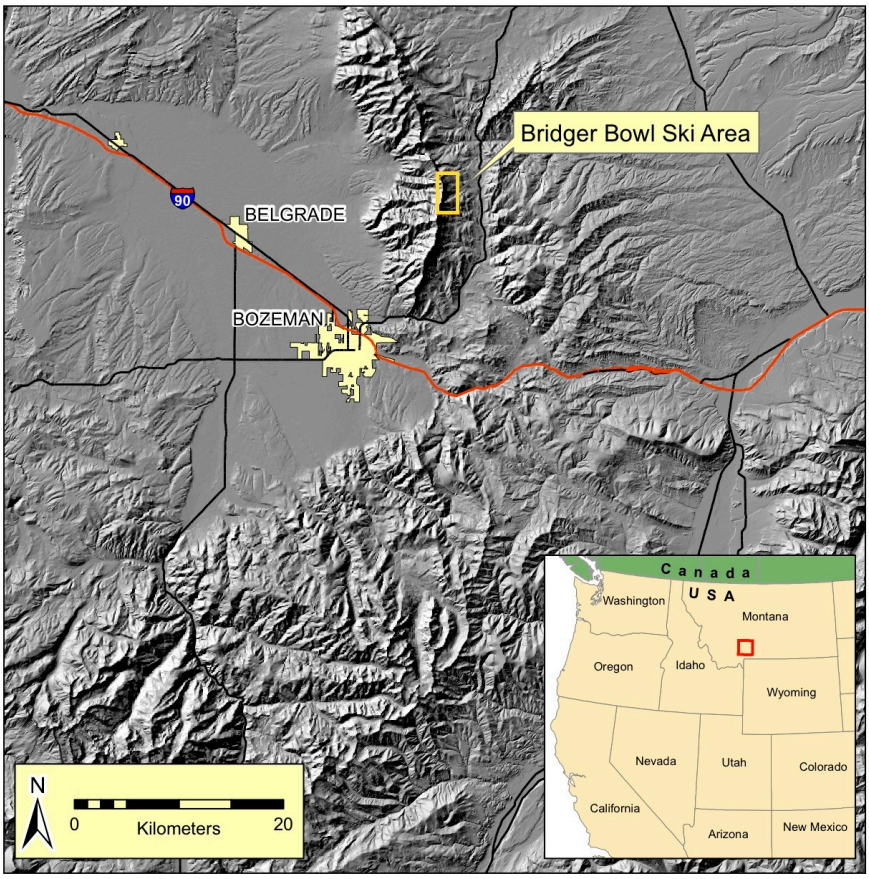



Figure 5. Location of Bridger Bowl ski area.

Bridger Bowl exhibits a seasonal intermountain climate with occasional winters showing continental or coastal climate conditions (Mock and Birkeland, 2000). The intermountain climate is distinguished from a coastal or coastal-transition climate by colder temperatures and less liquid precipitation. It is distinguished from continental climates by higher snowfall, warmer temperatures, and average temperature gradients above the threshold for constructive metamorphism (i.e., -10°C/m) in December (Mock and Birkeland, 2000).

Despite average December temperature gradients not meeting the threshold for constructive metamorphism, depth hoar is a common occurrence at the base of the

snowpack. Some years do have stronger temperature gradients, and frequent dry clear nights allow for outgoing longwave radiation to cool the snow surface and create a temperature gradient through the snowpack that is conducive to depth hoar formation. Furthermore, temperature gradients conducive to depth hoar formation may occur earlier than December, in October or November. Additionally, the snowpack structure at Bridger Bowl is dependent on the artificial alteration due to avalanche mitigation and frequent skier traffic.

Management of skier traffic in avalanche terrain has varied throughout Bridger Bowl's history. There has been an overall increase in skier visits since the mid 1980s (Fig. 6), but a more significant effect on the snowpack is the evolution of avalanche control, terrain use policies, and means of access. In order for public skiers to access the

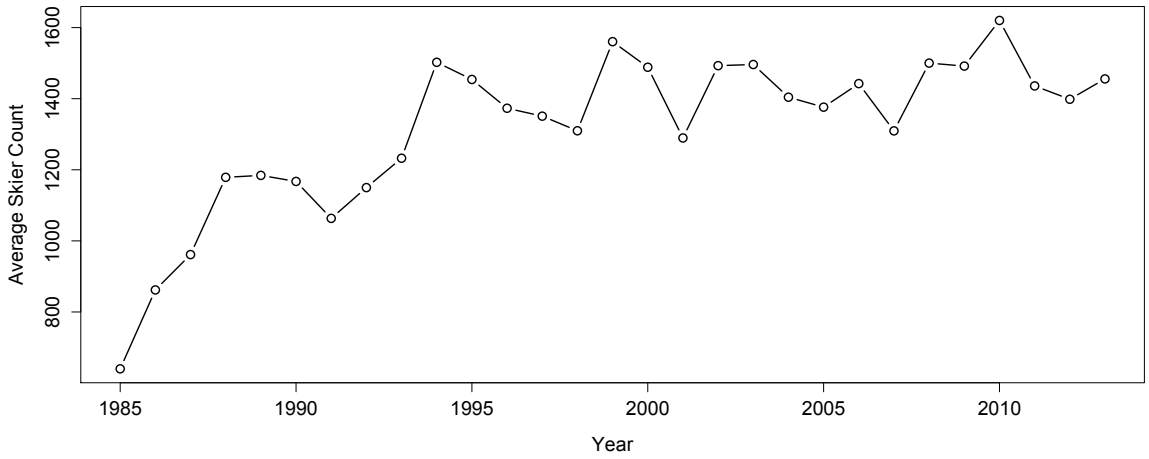



Figure 6. Annual average daily skier visits at Bridger Bowl since 1985.

majority of avalanche terrain at Bridger Bowl they have traditionally been required to carry an avalanche transceiver, shovel, and ski with a partner, in addition to hiking 500 vertical feet to the ridgeline. Recently, the Schlasman's chairlift was installed, which

provides lift access to mostly avalanche terrain and drops skiers off within a shorter distance from avalanche terrain that was previously harder to access. Skiers are now required to have only an avalanche transceiver in order to hike into avalanche terrain or ski terrain off the Schlasman's lift. In addition to easier and more widespread access to avalanche terrain, there has been an increased interest in riding in avalanche terrain driven by improvement in equipment and people's ability. These changes are important to consider, as the traffic patterns in avalanche starting zones will alter the snowpack structure and stability independently of weather. Skier visit data is available, but it is not sufficient to gauge the amount or location of travel in avalanche terrain. Therefore, it must be recognized that there has likely been an increase in skier compaction in avalanche terrain over time, which will affect the results of this study and limit the extrapolation of results to other areas.

Avalanche control methods at Bridger Bowl have evolved to include more reliance on the increase in skier traffic (Fig. 6) to help stabilize the snowpack. Avalanche terrain was traditionally mitigated using a fleet of 75mm recoilless rifles in combination with hand-deployed explosives. The use of the 75mm recoilless rifles became less and less desirable until the last shot was fired in the spring of 2008. The next year, the Schlasman's lift was installed, which provided effective access to use hand-deployed explosives on all avalanche terrain that required mitigation. This evolution has led to more timely and thorough avalanche mitigation, which allows for more skier traffic and theoretically quicker, more widespread stabilization of the snowpack.

## 3.2 Datasets

In this thesis we explore a case study of avalanches that fail on deeply buried persistent weak layers during the 2011-2012 season, and we perform a statistical analysis of meteorological metrics associated with these types of avalanches throughout recorded history (1968-2013) at Bridger Bowl. Meteorological data and avalanche control records were used for both the case study and the historical analysis. Bridger Bowl ski patrol has been recording weather and avalanche observations since 1968. Records of both exist from all seasons except for avalanche control records from the 1995-96 season, and weather records prior to February 1968. Due to the absence of these records, the analysis begins with the 1968-69 season, and the 1995-96 season is excluded. This yields a total of 44 seasons with meteorological and avalanche data that is sufficient for analysis between the 1968-69 and 2012-13 seasons. The duration of each season is considered to be from November to April, with the exact start and end date varying.

Avalanche and meteorological observations at Bridger Bowl followed the standard of the West Wide Avalanche Network (WWAN). Funding from the U.S. National Forest Service was cut from the WWAN program in 1995 (Williams, 1994), but recording standards have essentially remained the same. Recently, standards for recording meteorological and avalanche observations have evolved towards those set forth by Greene et al. (2010).

### 3.2.1 Meteorological Observations

Meteorological observations are recorded daily by ski patrol at the Alpine weather station (2286m) (Fig. 7). Of notable mention is the fact that snowfall increases with

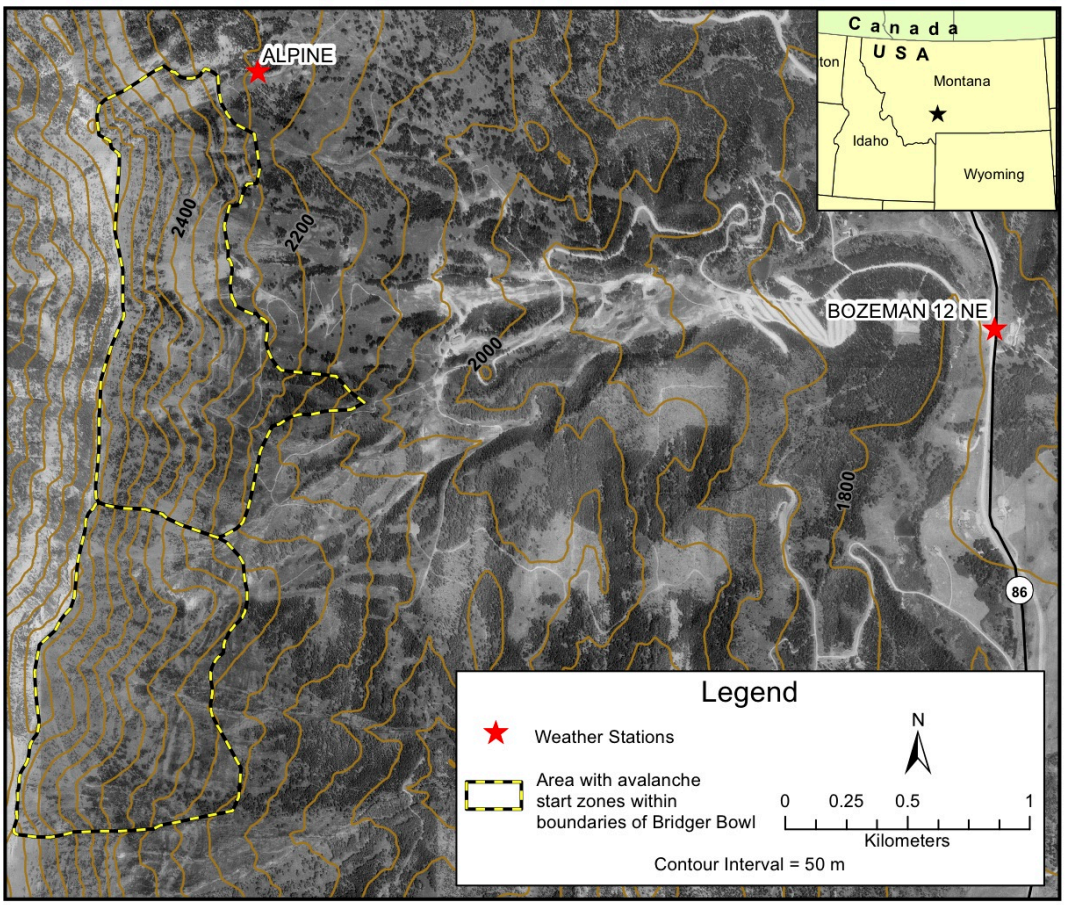



Figure 7. Locations of Alpine and Bozeman 12 NE COOP weather stations, and areas with avalanche starting zones at Bridger Bowl.

elevation and significant differences in storm totals are often observed between starting zones (~2550m) and the weather station (2286m), and temperature is typically lower in higher elevation starting zones. Meteorological observations were traditionally taken at 0800 hours, but were switched to 1600 hours in 2001. Snow depth, new snowfall, new snow water equivalent (SWE), and minimum and maximum temperature over a 24-hour period are included in the daily records (Table 4). These values were used to derive the metrics (section 3.4.3) used in this study that are believed to be associated with the formation and initiation of deep slabs on persistent weak layers. Precipitation and

temperature data were recorded in inches and Fahrenheit, respectively. Values were converted to metric (millimeters, centimeters and centigrade) before any analysis was performed.

Table 4 Meteorological observations from the Alpine weather station that are recorded daily by ski patrol at Bridger Bowl.

| Observed Daily Variables | | |
|---|---|---|
| Variable | Symbol | Description |
| Height of snow (cm) | HS | Height of snow on ground |
| 24 hr snowfall (cm) | HN | Height of new snow in last 24 h |
| Daily SWE (cm) | SWE | Total snow water equivalent (SWE) in 24 hours |
| Rain (mm) | Rn | Total rain over last 24 hours |
| Daily maximum temperature (°C) | Tmax | Maximum temperature over last 24 hours |
| Daily minimum temperature (°C) | Tmin | Minimum temperature over last 24 hours |

Meteorological data was missing from the months of November, December, and January in multiple years between 1968 and 1983 (Table 5). Meteorological data available through the Western Regional Climate Center (http://www.wrcc.dri.edu) were used from the nearby Bozeman 12 NE (COOP) weather station (Fig. 7) to estimate missing values at the Alpine weather station. Statistical relationships between the two stations were used to estimate values on days without measurements at the Alpine station from measured values at the COOP station. The dataset for the COOP station extends from June 1950 to December 1995. These data include snow depth, maximum temperature, minimum temperature, daily new SWE, and daily new snowfall. Data exist from both the Alpine and COOP stations for November, December, January, February, and March, from February 1968 to December 1995 with the exception of those months listed as absent in Table 5. Snow depth, maximum temperature, and minimum temperature from days that had data measured at both stations (Table 6) were used to find

statistical relationships between the two stations, and these relationships were used to estimate missing values at the Alpine station. Statistical methods used for this imputation are described in section 3.3.

Table 5 List of months that are missing from meteorological records at the Alpine weather station (N=November, D=December, J=January). Meteorological values in these months were estimated using statistical relationships between the Alpine and COOP weather stations.

| Season | Month(s) Missing |
|---|---|
| 1968 | N |
| 1969 | N |
| 1970 | N |
| 1971 | N |
| 1972 | ND |
| 1973 | N |
| 1974 | ND |
| 1975 | N |
| 1976 | ND |
| 1977 | N |
| 1978 | N |
| 1979 | NDJ |
| 1980 | NDJ |
| 1981 | N |
| 1982 | N |
| 1983 | N |

Table 6 Descriptive statistics for measured meteorological metrics at the Alpine weather station (2286m) and the COOP weather station (1813m). Summarized from all days that have a measurement from both stations.

| Alpine Metrics | Minimum | Maximum | Mean | Median |
|---|---|---|---|---|
| Snow depth (cm) | 0 | 320 | 149 | 147 |
| Maximum temperature (°C) | -34 | 20 | -1.1 | -1 |
| Minimum temperature (°C) | -42 | 7 | -9.5 | -9 |
| New snow over 24hrs. (cm) | 0 | 109.2 | 4.6 | 0 |
| Snow water equivalence over 24hrs. (mm) | 0 | 93 | 3.2 | 0 |
| **COOP Metrics** | **Minimum** | **Maximum** | **Mean** | **Median** |
| Snow depth (cm) | 0 | 130 | 43 | 43 |
| Maximum temperature (°C) | -31.1 | 20 | 2.3 | 2.8 |
| Minimum temperature (°C) | -43.9 | 8.9 | -11.5 | -10.6 |
| New snow over 24hrs. (cm) | 0 | 50 | 2.8 | 0 |
| Snow water equivalence over 24hrs. (mm) | 0 | 41 | 2 | 0 |

### 3.2.2 Avalanche Records

Historical avalanche records were used to identify deep slab avalanches that failed on persistent weak layers. All artificial and naturally triggered avalanches that occur within the boundaries of Bridger Bowl are recorded by ski patrol. Large avalanches that occur adjacent to the ski area are recorded, though not as thoroughly. Avalanche characteristics that are recorded with most observations and that are used in this study include: date, path name, type of trigger, avalanche type, relative size, crown depth, and layers involved (i.e., bed surface). The general guidelines used to record avalanche observations can be found in Greene et al. (2010). Some of these standards have changed since the termination of WWAN in 1995, so discrepancies in these data exist across years. Discrepancies that appear in the data due to changes in standards over the years, and how they were addressed in this thesis, are discussed below.

Recording standards for type of trigger, avalanche type, and relative size have few discrepancies as these have gone relatively unchanged (Greene et al., 2010). Crown depth is the average depth of the avalanche crown, which is the vertical face of the slab that remains intact with the slope at the top of an avalanche (Fig. 3). Crown depth was recorded in feet until the 1998-99 season, when observers began recording it in inches. All crown depths prior to the 1999-2000 season were multiplied by twelve in order to have all crown depths listed in inches. Observations during the 1998-99 season appeared to have some crown depths recorded in inches and some recorded in feet. Therefore, records from this season were examined closely to ensure that observations already recorded in inches were not converted.

Layers involved in an avalanche, also referred to as "avalanche bed surface" by Greene et al. (2010), has been recorded using a variety of codes over the years. The current standard (Greene et al., 2010) was not implemented until 2013 at Bridger Bowl, and therefore does not appear in the records used for this study. The avalanche records at Bridger Bowl contain a field for bed surface and a field for layers involved. Neither one of these fields has been recorded with every observation, so both were examined in detail to determine the layers involved in avalanches. These fields were important, in combination with crown depth, to define avalanches that failed on persistent weak layers since specific weak layer types are not noted in avalanche records. In the bed surface field, avalanches that ran on the ground were recorded with a "G", and avalanches that involved other layers were recorded with "O". In the layers field, "G" is also used for avalanches that ran on the ground, an "A" is used for an avalanche that involved only new snow, and "B" is used for an avalanche that involved old snow layers.

### 3.2.3 Snowpack Observations

Snowpack observations are not part of the data that is regularly recorded by the Bridger Bowl ski patrol. Therefore, weak layer types of historical avalanches are seldom known, so recorded avalanche characteristics were used to find avalanches that failed on deeply buried persistent weak layers for the historical data analysis. However, snowpack observations were recorded during the 2011-2012 season by the author, ski patrol, and the GNFAC. These observations will be used to support a case study from the 2011-2012 season.

Snowpack observations were collected during the 2011-2012 season in areas that were considered to have snowpack and terrain characteristics similar to avalanche starting zones within and adjacent to the ski area. These areas were revisited until the depth hoar layer appeared to be disturbed and no longer representative of less disturbed and likely weaker areas. Observers recorded stability test results, snowpack observations for each layer, and site characteristics (Greene et al., 2010). Temperature profiles were taken in the spring as the snow became isothermal. Average hardness of the slab was derived by averaging the hardness, weighted by thickness, of each layer above the weak layer. A stability index was created, which is based on whether the weak layer propagated during any extended column test (ECT) in that snowpit. Snowpits that had a propagating ECT were considered unstable and those that did not have a propagating ECT were considered stable.

## 3.3 Imputations for Missing Meteorological Data

Constructive metamorphism and the formation of persistent weak layers deep in the snowpack typically occur in November, December, and January when the snow is shallow and air temperatures are cold. Meteorological data at Bridger Bowl are absent for many of these months between 1968 and 1983 (Table 5), so missing values were estimated for these months using statistical relationships between weather observations at the Alpine weather station and the Bozeman 12 NE (COOP) weather station (Fig. 7).

### 3.3.1 Imputation Methods

#### 3.3.1.1 Temperature and Snow Depth Estimates

Minimum temperature, maximum temperature, and snow depth at the Alpine station were estimated based on their statistical relationship with the same variables measured at the COOP station. Relationships between variables were first explored graphically, and through fitting ordinary least squares (OLS) linear regression models and exploring plots of model residuals. Generalized additive models (GAMs) were used to attempt to fix any problems with linearity, normality, or heterogeneity in the OLS model residuals. GAMs were also used to explore models with different intercepts and relationships to the explanatory variable for individual months. Temporal correlation and trends were modeled using a smoothing function of the response variable over time. Models were fit that contained linear and smoother terms for explanatory variables, smoother terms for explanatory variables in individual months, terms for a different intercept for each month, and a temporal smoothing term (Zuur et al., 2006) using the "gam" function from the mgcv package in R statistical software (R Core Team, 2013). Smoothing functions are used to address issues of non-linear relationships. They are essentially a combination of different regression functions for each interval of a specified width along the x-axis. Polynomial (cubic) regressions are also used to create smoothing terms (Zuur et al., 2006 p. 38).

A subset of the data that included only the months of November, December, and January was used to model the three variables since we were only interested in estimating missing values in these months. Models were fit using a forward selection approach. First, the relationship between the variable being modeled at the Alpine station and the

same variable at the COOP station was modeled using both an OLS linear model and a GAM with a smoothing function for the explanatory variable. The more adequate of these was chosen based on having the lowest AIC scores (Zuur et al., 2006) and model residuals that better met assumptions. Next, temporal smoothing terms, terms used to create different intercepts for each month, and terms that allow smoothing functions for the explanatory variable in individual months, were added one by one. Lower AIC scores, improved $R^2$, model residuals more closely meeting assumptions, response v. fitted values being closer to a 1:1 relationship, and lower generalized cross validation (GCV) scores (Clark, 2013) were considered as strong evidence to suggest the inclusion of the added term. GCV scores are calculated using a leave one out cross validation process, and can be considered as a similar estimate to a mean square prediction error (Clark, 2013). In addition to these criteria, p-values from ANOVA F-tests were considered as strong evidence to include model terms. P-values in these types of models are only approximate, and should not be used with a strict cut off (Clark, 2013; Zuur et al., 2006). Therefore, p-values were only considered as strong evidence for inclusion of a term if they were very low (lower than 0.01).

3.3.1.2 Snow Depth Estimate Corrections A correction was applied to the estimated snow depth values in each season in order to ensure continuity in accumulation trends. This correction was based on the difference between each season's last estimated snow depth value and its first measured snow depth value. The approximate difference between each season's first measured snow depth and the sum of the last modeled snow

depth plus the first observed measured 24-hour snowfall was added to all modeled snow depth values for that season. This can be written as:

$HS_x$(estimated) = $HS_x$(modeled) + $HS_{i+1}$(measured) – ($HS_i$(modeled) + $HN_{i+1}$(measured)),

where HS is snow depth, i is the last day of estimated values in season x, and HN is 24 hr snowfall. For example, in the 1977 season the last modeled value was approximately 15cm less than the first measured value at Alpine and there was no measured new snow on the first day of observation at Alpine. Therefore, 15cm was added to all modeled values for that season (Fig. 8). In 1980 there was approximately no difference between the first measured snow depth at Alpine and the last modeled snow depth. However, there was 20cm of new snow measured on the first observed day at Alpine, so 20cm was subtracted from all modeled values for that season (Fig. 8).

3.3.1.3 New Snowfall and Snow Water Equivalent Estimates New snowfall and snow water equivalent amounts graphically showed weak evidence of a linear relationship between the two stations (Fig. 9). Therefore, missing values for new snowfall at Alpine were estimated using positive daily changes in estimated snow depth rather than a statistical relationship with the COOP station's new snowfall values. Snow water equivalent was estimated by multiplying the estimated new snowfall amount by a new snow density that was estimated from maximum temperature based on an OLS regression equation fit from historical values. Historical density values were calculated as %H2O using historical measured new snowfall heights and new SWE in the equation:

% H20 = HN/SWE,

where HN is the height of new snow and SWE is the new snow water equivalent.

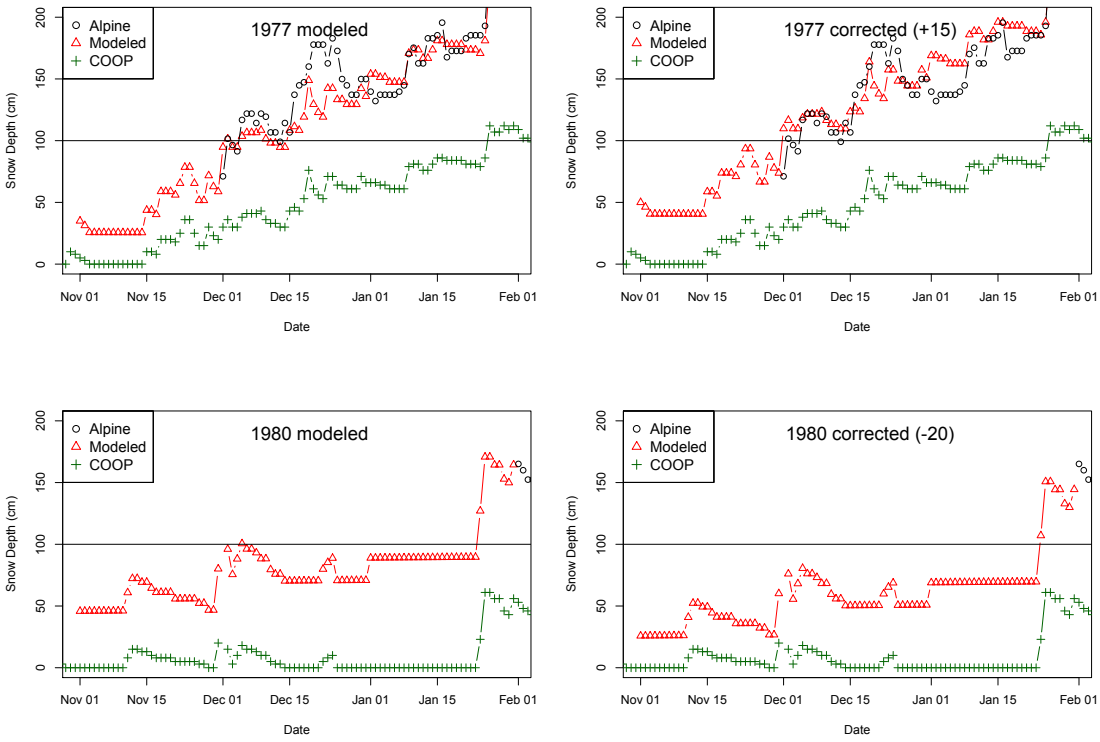



Figure 8. Time series plots of modeled Alpine meteorological values and measured values from both stations, shown for two seasons (1977-78 and 1980-81) before and after applying a correction to modeled values.

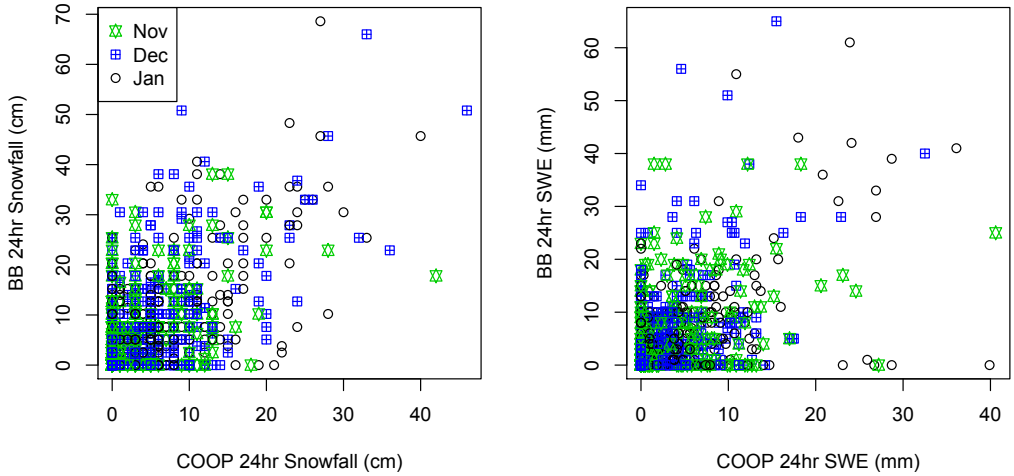



Figure 9. Plots of new snowfall and new snow water equivalent (SWE) at the Bridger Bowl (BB) weather station versus those metrics at the COOP weather station.

3.3.2 Imputation Results

All three variables (maximum temperature, minimum temperature, and snow depth) show generally strong linear relationships with their respective variables at the other weather station (Fig. 10). There is some graphical evidence to suggest different ranges of values and different relationships in different months for temperature variables, and strong graphical evidence for different intercepts and relationships in different months for snow depth (Fig. 10). Therefore, terms were tested that may represent the various possible different relationships. A smoothing term for the explanatory variable was shown to be more adequate than a linear term for all three variables, based on lower AIC scores (Table 7) and residuals more closely meeting assumptions of normality and linearity for the model with the smoothing term (Appendix A). A term to include separate intercepts for each month was shown to be significant for all three variables based on strong evidence from ANOVA F-tests, lowered AIC scores, lower GCV scores, and improved $R^2$ when the monthly intercept term (factor(month)) was added to the model (Table 7). A temporal smoothing function (s(time)) was also shown to be significant in explaining a portion of the variance of all three variables based on strong evidence from the same criteria as above when added to the model (Table 7).

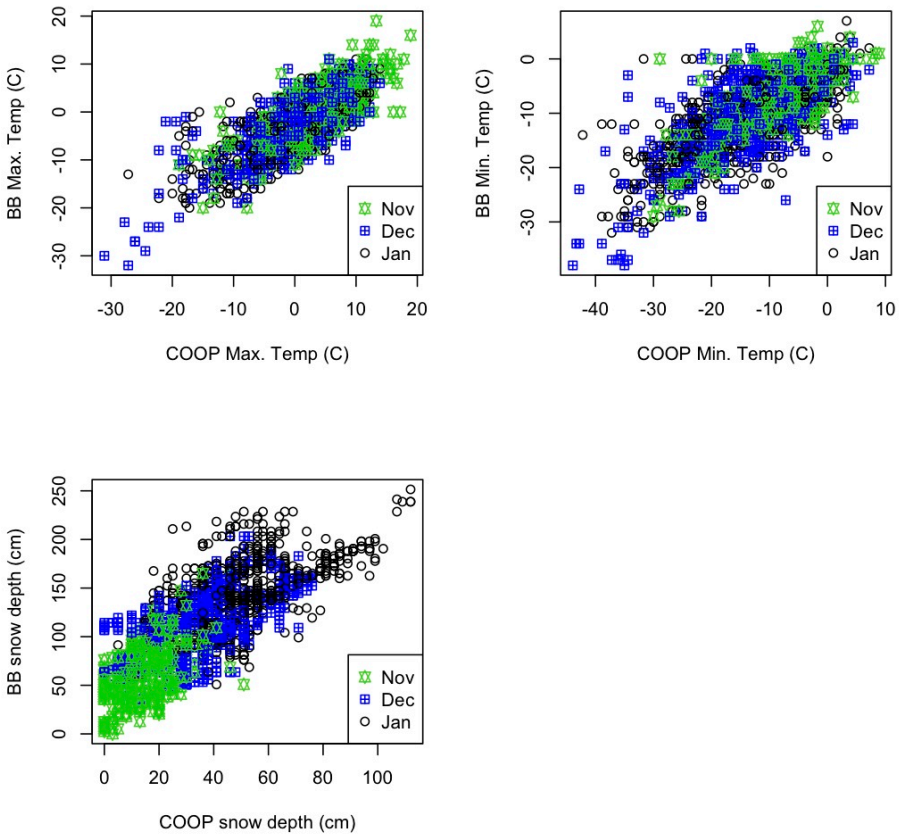



Figure 10. Plots of temperature and snow depth variables at Bridger Bowl (BB) versus the COOP station.

Table 7 AIC scores and ANOVA test results from model selection for temperature and snow depth from both datasets. Blank cells in the middle columns indicate unobtainable or not applicable test results. Blank cells in the AIC, GCV, or $R^2$ column indicate that row is an ANOVA test on a different term in the above model.

| Maximum Temperature | | | | | | | |
|---|---|---|---|---|---|---|---|
| Current Model | F-Tested Term | F-stat | p-value | AIC | GCV | edf | $R^2$ |
| ~Ntmx | | | | 10123.80 | 14.8 | | 60.0 |
| ~s(Ntmx) | | | | 10089.26 | 14.5 | 8.0 | 61.1 |
| ~s(Ntmx)+factor(month) | factor(month) | 7.92 | 0.0003 | 10077.10 | 14.5 | | 61.4 |
| ~s(Ntmx)+factor(month)+s(time) | s(time) | 11.52 | 0.0000 | 9993.41 | 13.8 | 8.8 | 63.5 |
| ~s(time)+s(Ntmx)+factor(month)+s(Nov.) | s(Nov.) | 0.56 | 0.4521 | 9994.82 | 13.8 | 1.5 | 63.5 |
| ~s(time)+s(Ntmx)+factor(month)+s(Dec.) | s(Dec.) | 5.87 | 0.0003 | 9979.93 | 13.7 | 8.6 | 63.9 |
| ~s(time)+s(Ntmx)+factor(month)+s(Jan.) | s(Jan.) | 3.24 | 0.0019 | 9984.40 | 13.7 | 7.1 | 64.0 |
| ~s(time)+s(Ntmx)+factor(month)+s(Jan.)+s(Dec.) | s(Dec.) | 0.72 | 0.5525 | 9980.29 | 13.7 | 8.6 | 63.9 |
| | s(Jan.) | 1.15 | 0.2238 | | | 3.6 | |
| ~s(time)+s(Ntmx)+factor(month)+s(Dec.)+s(Nov.) | s(Nov.) | 1.86 | 0.1660 | 9980.11 | 13.7 | 2.2 | 64.0 |
| | s(Dec.) | 5.63 | 0.0002 | | | 8.6 | |
| ~s(time)+s(Ntmx)+factor(month)+s(Jan.)+s(Nov.) | s(Nov.) | | | 9986.39 | 13.8 | 1.4 | 64.0 |
| | s(Jan.) | 3.16 | 0.0024 | | | 7.1 | |
| ~s(time)+s(Ntmx)+factor(month)+s(Dec.)+s(Nov.)+s(Jan.) | s(Nov.) | | | 9979.09 | 13.7 | 1.2 | 63.9 |
| | s(Dec.) | 0.48 | 0.7873 | | | 7.8 | |
| | s(Jan.) | 1.38 | 0.2519 | | | 3.3 | |
| Minimum Temperature | | | | | | | |
| Current Model | F-Tested Term | F-stat | p-value | AIC | GCV | edf | R^2 |
| ~Ntmn | | | | 10710.51 | 20.4 | | 53.3 |
| ~s(Ntmn) | | | | 10661.16 | 19.9 | 5.5 | 54.8 |
| ~s(Ntmn)+factor(month) | factor(month) | 10.13 | 0.0000 | 10644.35 | 19.7 | | 55.3 |
| ~s(time)+s(Ntmn)+factor(month) | s(time) | 14.06 | 0.0000 | 10541.81 | 18.6 | 8.7 | 58.1 |
| ~s(time)+s(Ntmn)+factor(month)+s(Nov.) | s(Nov.) | 8.60 | 0.0001 | 10526.84 | 18.5 | 2.7 | 58.6 |
| ~s(time)+s(Ntmn)+factor(month)+s(Dec.) | s(Dec.) | 4.13 | 0.0004 | 10528.78 | 18.5 | 6.5 | 58.7 |
| ~s(time)+s(Ntmn)+factor(month)+s(Jan.) | s(Jan.) | 4.42 | 0.0002 | 10526.63 | 18.5 | 6.6 | 58.7 |
| ~s(time)+s(Ntmn)+factor(month)+s(Nov.)+s(Dec.) | s(Dec.) | 3.76 | 0.0006 | 10514.80 | 18.4 | 6.5 | 59.1 |
| | s(Nov.) | 6.73 | 0.0002 | | | 3.0 | |
| ~s(time)+s(Ntmn)+factor(month)+s(Nov.)+s(Jan.) | s(Jan.) | 3.91 | 0.0005 | 10514.62 | 18.4 | 6.8 | 59.1 |
| | s(Nov.) | 7.10 | 0.0004 | | | 2.2 | |
| ~s(time)+s(Ntmn)+factor(month)+s(Jan.)+s(Dec.) | s(Dec.) | 8.10 | 0.0002 | 10513.43 | 18.3 | 6.0 | 59.1 |
| | s(Jan.) | 8.63 | 0.0001 | | 5.2 | 5.2 | |
| ~s(time)+s(Ntmn)+factor(month)+s(Nov.)+s(Dec.)+s(Jan.) | s(Jan.) | | | 10514.28 | 18.4 | 6.3 | 59.1 |
| | s(Dec.) | 3.10 | 0.0963 | | | 6.7 | |
| | s(Nov.) | 0.26 | 0.4465 | | | 3.0 | |

Table 7 (Continued)

| Height of Snow Current Model | F-Tested Term | F-stat | p-value | AIC | GCV | edf | R^2 |
|---|---|---|---|---|---|---|---|
| ~Nhs | | | | 17165.98 | 887.7 | | 58.5 |
| ~s(Nhs) | | | | 17104.91 | 857.8 | 5.6 | 60.1 |
| ~s(Nhs)+factor(month) | factor(month) | 117.61 | 0.0000 | 16571.45 | 731.1 | | 66.1 |
| ~s(Nhs)+s(time)+factor(month) | s(time) | 93.69 | 0.0000 | 16213.70 | 520.4 | 9.0 | 76.1 |
| ~s(Nhs)+s(time)+factor(month)+s(Nov.) | s(Nov.) | | | 16212.76 | 520.1 | 1.5 | 76.1 |
| ~s(Nhs)+s(time)+factor(month)+s(Dec.) | s(Dec.) | 2.30 | 0.0993 | 16213.04 | 520.2 | 2.7 | 76.1 |
| ~s(Nhs)+s(time)+factor(month)+s(Jan.) | s(Jan.) | 2.66 | 0.0153 | 16209.73 | 519.3 | 6.9 | 76.3 |
| ~s(Nhs)+s(time)+factor(month)+s(Jan.)+ s(Dec.) | s(Dec.) | 0.70 | 0.5517 | 16205.85 | 518.1 | 1.4 | 76.3 |
| | s(Jan.) | 11.10 | 0.0021 | | | 6.7 | |
| ~s(Nhs)+s(time)+factor(month)+s(Jan.)+ s(Nov.) | s(Nov.) | | | 16199.05 | 516.2 | 4.3 | 76.4 |
| | s(Jan.) | 4.27 | 0.0003 | | | 6.8 | |
| ~s(Nhs)+s(time)+factor(month)+s(Nov.) +s(Dec.) | s(Dec.) | 2.65 | 0.0037 | 16206.26 | 518.3 | 5.0 | 76.4 |
| | s(Nov.) | 2.88 | 0.0043 | | | 4.5 | |
| ~s(Nhs)+s(time)+factor(month)+s(Jan.)+ s(Nov.)+s(Dec.) | s(Dec.) | | | 16199.05 | 516.2 | 1.2 | 76.4 |
| | s(Nov.) | 4.27 | 0.0055 | | | 4.1 | |
| | s(Jan.) | 0.08 | 0.9866 | | | 6.5 | |

3.3.2.1 Maximum Temperature Adding a separate smoothing function for the month of December was shown to be significant in the model for maximum temperature based on a low p-value from the ANOVA F-test, improved $R^2$, lower AIC, and a large estimated degrees of freedom (edf) for the smoothing term (Table 7). The edf for a smoothing term is an estimate of the smoothers wiggliness (Zuur et al., 2006). An edf close to one indicates a linear relationship. There was also strong evidence to include a smoothing term for January based on lower AIC and improved $R^2$. However, the p-value is not as small and is less convincing than that for the December smoothing term. There was a lack of evidence to include a smoothing function for November or January after including a smoothing function for December (Table 7). Therefore, the most adequate model for maximum temperature was chosen as:

BBmaxT = s(NmaxT) + s(time) + factor(month) + s(NmaxT, month=December),

where BBmaxT is the maximum temperature at the Alpine weather station, NmaxT is the maximum temperature at the COOP station, s(time) is a temporal smoothing function, factor(month) is a term to include a separate intercept for each month, and s(NmaxT, month=December) is a term to include a smoothing function for maximum temperature in December. Residuals from this model show improvement in normality and linearity over those from the model with only one linear or smoothing term (Appendix A). These residuals and the improved $R^2$ and GCV scores over simpler models support the adequacy of this model for estimating maximum temperature at the Alpine weather station.

3.3.2.2 Minimum Temperature There is strong evidence to include separate smoothing functions for the months of November, December, and January in the model for minimum temperature if a smoothing function is only included for one or two of those months (Table 7). The model with a smoothing function for January and December has the lowest AIC score making it the choice as the most adequate model for minimum temperature from the November-January dataset. The full form of this model is:

BBminT = s(NminT) + s(time) + factor(month) + s(NminT, month=December) + s(NminT, month=January),

Which uses similar notation as above models.

3.3.2.3 Snow Depth The most adequate model for snow depth based on AIC scores, GCV scores, and lowest $R^2$ was:

BBhs = s(Nhs) + s(time) + factor(month) + s(Nhs, month=November) + s(Nhs, month=January).

However, the ANOVA test for the November smoothing term failed to provide a result, and neither of the November or January smoothing terms are significant before the other one is added. The January term has a relatively large p-value before accounting for other monthly smoothing terms, and the November term has no ANOVA result and an edf of 1.5, which indicates little need for a non-linear term, before adding other monthly smoothing terms. Therefore, a model without any monthly smoothing terms was used to estimate snow depth, in the form of:

BBhs = s(Nhs) + s(time) + factor(month).

Estimated values from this model were graphed against measured values from both weather stations for each season (e.g., Fig. 8), and a correction was applied to each seasons estimated values as described above (Table 8).

The mean estimated snow depth at the Alpine weather station was 100.0cm with an associated 95% confidence interval from 95.2cm to 104.9cm. The maximum estimated snow depth was 229.9cm with an associated 95% confidence interval from 215.0cm to 229.9cm. The Minimum estimated snow depth was 0.4cm with an associated 95% confidence interval from -9.5cm to 10.4cm. The range of confidence intervals for estimated snow depth ranged from +/- 9.7cm to +/-19.9cm.

Table 8 Corrections applied to estimated snow depth values, and which months had values estimated in each season at the Alpine weather station.

| Season | Month(s) | Correction |
|---|---|---|
| 1968 | N | 0 |
| 1969 | N | 20 |
| 1970 | N | 0 |
| 1971 | N | -30 |
| 1972 | ND | -25 |
| 1973 | N | 40 |
| 1974 | ND | -20 |
| 1975 | N | -10 |
| 1976 | ND | -40 |
| 1977 | N | 15 |
| 1978 | N | 5 |
| 1979 | NDJ | -40 |
| 1980 | NDJ | -20 |
| 1981 | N | -20 |
| 1982 | N | -25 |
| 1983 | N | -10 |

3.3.2.4 New Snowfall and Snow Water Equivalent Twenty-four hour snowfall amounts were estimated from positive changes in modeled snow depth. A relationship between measured maximum temperature and %H20 was used to estimate new snow density, which was multiplied by the estimated new snowfall amount to estimate new SWE.

The linear fit between maximum 24-hour temperature and %H20 improved with a natural log transformation on the response (Fig. 11). The equation found through OLS regression, and used to estimate SWE, was:

$\text{Log}(\%H2O) = -2.628091 + 0.041439 * Tmax,$

with temperature in degrees Celsius.

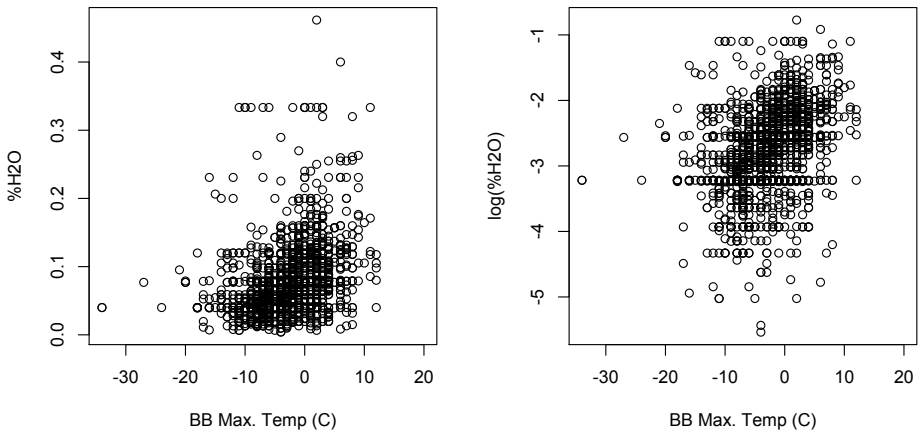



Figure 11. Relationship between new snow density and maximum temperature. A natural logarithmic transformation of density was used to fit an OLS linear regression with maximum temperature (right).

The equation for percent water of new snow provides a new snow density of 72 kg/m$^3$ when the maximum temperature is 0°C. This is similar to the average (80 kg/m$^3$) and median (66 kg/m$^3$) density of recorded new snowfall at Bridger Bowl before 1995 (Table 9). This is also the same as the lower end of average new snow density values recorded by Judson and Doesken (1999) at six sites in Colorado and Wyoming over four years (Table 9). LaChapelle (1962) also found similar average new snow density values at areas with similar climates (intermountain-continental) as Bridger Bowl (Table 9). The average maximum temperature on days with new snow at Bridger Bowl was -2.7°C. At this temperature the equation gives a new snow density of 65 kg/m$^3$. This is similar to the median new snow density (66 kg/m$^3$) recorded at Bridger Bowl before 1995 as well as mean and median new snow density values reported in similar climates (Table 9). The similarity of estimated new snow density values to previously observed new snow density

values supports the use of the equation found in this thesis to estimate missing new SWE values at Bridger Bowl.

Table 9 Summary of snow density recorded at Bridger Bowl before 1995, at six sites in Colorado and Wyoming (Judson and Doesken, 1999), and at one site in five different states (LaChapelle, 1962). The range of means and medians is shown for the Colorado and Wyoming sites studied by Judson and Doesken (1999).

| Location | New snow density (kg/m^3) | | | | Climate Classification |
|---|---|---|---|---|---|
| | Min. | Max. | Mean | Median | (Mock and Birkeland, 2000) |
| Bridger Bowl | 4 | 462 | 80 | 66 | Intermountain |
| 6 sites in CO. and WY. (Judson and Doesken, 1999) | 10 | 257 | 72 - 103 | 70 - 100 | Continental/Intermountain |
| Alta, UT (Lachapelle, 1962) | NA | NA | 100 | 90 | Intermountain/Coastal |
| Stevens Pass, WA (LaChapelle, 1962) | NA | NA | 90 | 80 | Coastal |
| Berthoud Pass, CO (LaChapelle, 1962) | NA | NA | 70 | 65 | Continental |
| Squaw Valley, CA (LaChapelle, 1962) | NA | NA | 90 | 90 | Coastal |
| Girdwood, AK (LaChapelle, 1962) | NA | NA | 70 | 70 | Coastal/Continental |

### 3.3.3 Imputation Conclusion

Statistical relationships of observed meteorological metrics were found between the Bridger Bowl and COOP weather stations. These relationships were then used to estimate missing values for these metrics at Bridger Bowl in the months of November, December, and January from 1968-1983 (Table 5). The estimated and observed values at Bridger Bowl were then used to derive meteorological metrics that may be associated with deep slab avalanches on persistent weak layers.

## 3.4 Historical Analysis of Avalanches at Bridger Bowl 1968-2013

Deep slab avalanches that failed on persistent weak layers late in the season at Bridger Bowl were derived from historical avalanche records and correlated with observed and estimated historical meteorological metrics. Seasons and days that had one

or more avalanches that failed on a deep persistent weak layer in February, March, or April were examined for associations with meteorological metrics that are believed to be contributors to the formation and initiation of avalanches that fail on deep persistent weak layers. Univariate analysis, classification trees (Breiman et al., 1993), and random forests (Breiman, 2001) were used to find meteorological metrics that best discriminated between seasons with deep slabs on persistent weak layers and seasons without deep slabs on persistent weak layers. These methods were also used to find metrics that discriminate between days with and days without deep slabs on persistent weak layers, when a deep persistent weak layer was likely present.

Previous studies have used similar methods that employ classification trees to examine differences in meteorological metrics between days with and without avalanches (e.g., Davis et al., 1999; Hendrikx et al., 2005; Peitzsch et al., 2012). These studies generally focus on any avalanche occurrence, or a threshold number of avalanches, on a given day. They also primarily focus on meteorological metrics up to around seven days prior to days with avalanche activity. This thesis examines meteorological metrics up to seven days prior to days that had deep slabs on persistent weak layers, as well as metrics that summarize meteorological conditions over the entire season and the early months of the season, when basal persistent weak layers are formed.

### 3.4.1 Deep Slabs on Persistent Weak Layers

Weak layer type was not recorded with avalanche observations, so other characteristics that are available in the historical records (section 3.2) were used to imply whether or not an avalanche failed on a deeply buried persistent weak layer. All

avalanches recorded as sliding on the ground were assumed to have failed on depth hoar, because that is a common weak layer in avalanches that slide on the ground in this climate. However, not all avalanches that fail on depth hoar are recorded as sliding on the ground. This is partly due to different interpretations between observers of what "sliding on the ground" means. Depth hoar avalanches often collapse along the upper boundary of the thick depth hoar layer (Bradley et al., 1977), resulting in depth hoar grains below this interface being left in place on the slope. Some consider this to mean the avalanche slid on the depth hoar and did not slide on the ground, and therefore record it as sliding on old layers. Others consider this type of avalanche to have slid on the ground because the weak layer is on the ground, and record it as sliding on the ground. Another situation that may prevent an observer from recording an avalanche that failed on depth hoar as an avalanche on the ground would be one where a thin melt-freeze layer lies below the depth hoar, meaning the depth hoar is not the lowest layer. Furthermore, avalanches that failed on layers of facets near the ground would not be recorded as sliding on the ground, but are important to consider for this study.

If avalanches after February 1st were not recorded as sliding on the ground, the relative size, avalanche type, and crown depth were used to see if they appeared to fail on a deep persistent weak layer. First, avalanches with R-sizes greater than or equal to two and crown depths greater than 36 inches (0.9m) were selected. Avalanches were chosen from this selection if they appeared larger in comparison with other events during the same cycle, occurred without the addition of new snow, were significantly deeper than preceding days storm totals, or involved old snow layers and had similar crown depths to avalanches that were recorded as sliding on the ground during the same or adjacent

cycles. Each event was examined individually to see that it met criteria suggesting that it failed on a persistent weak layer. Avalanches that were implied to fail on a persistent weak layer late in the season, and were not recorded as sliding to the ground, were required to have an avalanche type of hard slab or wet slab or similar characteristics to avalanches that were recorded as sliding on the ground during the same or adjacent cycle. Soft slabs and wet loose avalanches were included if the layers involved, or bed surface, was recorded as the ground. Therefore, wet loose avalanches that failed on persistent weak layers are included in addition to all slab avalanches (i.e., wet slab, soft slab, or hard slab) that failed on persistent weak layers. Avalanches from 2010-2013 had weak layers known to the author, so avalanches failing on deep persistent weak layers for these seasons were defined based on the known weak layer.

Avalanches failed on deep persistent weak layers after February 1st during 24 of the 44 seasons studied (Table 10). There were a total of 196 events on 76 days in February, March and April. Of these days, 50 had dry slab avalanches and 26 had wet snow avalanches. Five seasons had only wet snow avalanches, seven seasons had only dry slab avalanches, and the other twelve seasons had both wet and dry snow avalanches that failed on deep persistent weak layers after February 1$^{st}$.

Table 10 Summary of the 24 seasons that had at least one avalanche failing on a deep persistent weak layer late in the season. Seasons are defined by the year in which they started (e.g., 1968 was from November 1968-April 1969).

| Season | # of days | Wet snow days | # of slides | Wet snow slides | HS | SS | WL | WS |
|---|---|---|---|---|---|---|---|---|
| 1968 | 3 | 1 | 8 | 4 | 3 | 1 | 0 | 4 |
| 1971 | 2 | 1 | 3 | 1 | 2 | 0 | 0 | 1 |
| 1974 | 2 | 0 | 2 | 0 | 0 | 2 | 0 | 0 |
| 1975 | 5 | 4 | 49 | 48 | 0 | 1 | 10 | 38 |
| 1976 | 2 | 1 | 2 | 1 | 1 | 0 | 0 | 1 |
| 1977 | 1 | 1 | 4 | 4 | 0 | 0 | 4 | 0 |
| 1978 | 2 | 1 | 2 | 1 | 1 | 0 | 1 | 0 |
| 1979 | 3 | 0 | 3 | 0 | 3 | 0 | 0 | 0 |
| 1980 | 3 | 0 | 13 | 0 | 3 | 10 | 0 | 0 |
| 1981 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 1984 | 6 | 3 | 22 | 12 | 9 | 1 | 0 | 12 |
| 1985 | 3 | 1 | 6 | 2 | 4 | 0 | 0 | 2 |
| 1986 | 4 | 1 | 11 | 7 | 0 | 4 | 0 | 7 |
| 1987 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 |
| 1989 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| 1991 | 1 | 1 | 2 | 2 | 0 | 0 | 2 | 0 |
| 1992 | 4 | 2 | 9 | 5 | 3 | 1 | 3 | 2 |
| 1993 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| 2000 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 |
| 2002 | 9 | 1 | 13 | 2 | 3 | 8 | 2 | 0 |
| 2003 | 2 | 1 | 3 | 2 | 1 | 0 | 2 | 0 |
| 2006 | 4 | 0 | 10 | 0 | 1 | 9 | 0 | 0 |
| 2009 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| 2011 | 14 | 4 | 29 | 17 | 10 | 2 | 0 | 17 |
| Total | 76 | 26 | 197 | 111 | 46 | 40 | 26 | 85 |

### 3.4.2 Days and Seasons without Deep Slabs

It is typically desirable to have an almost equal number of events as non-events for statistical comparisons in order to avoid skewed distributions of data (Bois et al., 1975). While a benefit of classification trees is that they are insensitive to underlying distributions of data (Breiman et al., 1993), having an equal number of non-events during similar times as events is beneficial to control for temporal variations in site specific factors (e.g., climate and snowpack). Including events and non-events from similar times

is more practical for forecasting purposes. We are interested in discriminating between days with events and days with non-events during the time of the season when we are forecasting deep slabs, not at the beginning of the season or during seasons without deep persistent weak layers.

All seasons without deep slabs on persistent weak layers were compared to seasons with deep slab on persistent weak layers. There were 24 seasons with deep slabs on persistent weak layers and 20 seasons without deep slabs on persistent weak layers (Table 10).

Days without deep slabs on persistent weak layers were chosen so that there were an equal number of days without deep slabs on persistent weak layers as days with deep slabs on persistent weak layers in each season. Days without avalanches of this type were randomly selected from days that had avalanche control in February, March, or April, and were before the date of the last deep slab on a persistent weak layer. These criteria allow for assessment of the difference between meteorological metrics on days that explosive triggers were applied to the snowpack with and without triggering deep slabs on persistent weak layers, when a deep persistent weak layer was present and potentially active.

### 3.4.3 Meteorological Metrics

Daily and seasonal meteorological metrics that are believed to be associated with the formation and initiation of deep slabs on persistent weak layers were defined as explanatory variables. Summaries of meteorological metrics on days prior to avalanche activity are referred to as “daily” metrics, and summaries of meteorological metrics from

the entire season and early months of the season are referred to as "seasonal" metrics. Seasonal metrics were compared between seasons with and seasons without deep slabs on persistent weak layers. Seasonal and daily metrics were compared between days with deep slabs on persistent weak layers and days without these avalanches that still had deep persistent weak layers present and other avalanche activity.

3.4.3.1 Daily Meteorological Metrics Observed daily meteorological metrics were used (Table 4) alongside metrics that summarize meteorological conditions over several days prior to an event (Table 11), which were derived from observed metrics. Total snowfall and SWE over 2,3,5, and 7 days was derived because cumulative load, and delayed response to loading (up to 19 days) has been shown to be associated with deep slabs on persistent weak layers (e.g., Jamieson et al., 2001; Savage, 2006). Maximum and average temperatures (both maximum and minimum) over 2,3,5, and 7 days were derived to see if magnitude or trend of prior temperature is associated with avalanche initiation on deep persistent weak layers. The difference in air temperature (both minimum and maximum) between the current day and previous day was used to test for an association with large changes in temperature over 24 hours since rapid change is suggested to contribute to avalanche initiation (Perla, 1970; Tracz, 2012).

The total degrees above zero Celsius for the minimum and maximum temperature over 1, 2, 3, and 5 days was derived to assess the relationship between prolonged warm (above freezing) temperatures and deep slabs on persistent weak layers. This metric was primarily used to test for an association with triggering mechanisms related to wet avalanches because over 50% of avalanches in this study were wet avalanches, and 33%

of days with deep slabs on persistent weak layers had wet avalanches (Table 10). Additionally, total positive air temperature over 1, 3, and 5 days has previously been found to be associated with wet slab and glide avalanches (Peitszch et al., 2012). The change in snow depth over 2, 3, 5, and 7 days was also used, primarily as an indicator of settlement, which has also been found to be associated with wet slab and glide avalanches (Peitzsch et al., 2012). This variable is suggested as an indicator of snowpack consolidation and rate of water infiltration due to melting snow (Peitzsch et al., 2012).

Table 11 Daily meteorological metrics that were derived to compare conditions that precede days with and days without deep slabs on persistent weak layers.

| Derived Daily Variables<br>Variable | Symbol | Time |
|---|---|---|
| Total (sum) 24 hr snowfall (cm) | HNi | 2,3,5, and 7 days |
| Total (sum) daily SWE (cm) | SWEi | 2,3,5, and 7 days |
| Change in snow depth (cm) | HSi | 2,3,5, and 7 days |
| Max. air temperature during time i (°C) | maxTi | 2,3,5, and 7 days |
| Min. air temperature during time i (°C) | minTi | 2,3,5, and 7 days |
| Average min. air temperature (°C) | Tminavi | 2,3,5, and 7 days |
| Average max. air temperature (°C) | Tmaxavi | 2,3,5, and 7 days |
| Degrees above zero of min. temperature (°C) | minAbovezi | 1,2,3, and 5 days |
| Degrees above zero of max. temperature (°C) | maxAbovezi | 1,2,3, and 5 days |
| Difference in minimum temperature (°C) | TminDif2 | between current day and previous |
| Difference in maximum temperature (°C) | TmaxDif2 | between current day and previous |

3.4.3.2 Seasonal Meteorological Metrics Meteorological metrics that summarize processes related to the formation and initiation of deep slabs on persistent weak layers were derived from daily meteorological records for all 44 seasons studied. These were used as explanatory variables in the historical analysis of deep slabs on persistent weak layers and were summarized for the case study of the 2011-2012 season. Seasonal meteorological metrics are summaries of temperature, daily new snowfall, daily

new SWE, daily snow depth, and calculated daily temperature gradients (Table 12). Daily Temperature gradients were calculated by dividing the daily minimum temperature by the snow depth. Historical daily values for each metric were summarized for each season over the entire season, individual months (November-March), and aggregations of the months November and December, and November, December and January (Table 12).

Table 12 Seasonal meteorological metrics that were derived to compare conditions during seasons with and seasons without deep slabs on persistent weak layers.

| Derived Seasonal Variables | | | |
|---|---|---|---|
| Variable | Symbol | Summaries | Time |
| Height of snow (cm) | HS | avg, max | Nov, Dec, Jan, Season |
| Daily SWE (cm) | SWE | total (sum) | Nov, Dec, Jan, ND, NDJ |
| | | avg, max, sum | Season |
| 24 hr snowfall (cm) | HN | total (sum) | Nov, Dec, Jan, ND, NDJ |
| | | avg, max, sum | Season |
| Daily temperature gradient (°C/10cm) | tg | avg | Nov, Dec, Jan, Feb, Mar, ND, NDJ |
| | TG | avg, max, min | Season |
| Maximum temperature (°C) | Tmx | max, min | Season |
| Minimum temperature (°C) | Tmn | max, min | Season |
| Days with TG>10°C/m & HS<1m | tgcnt | proportion | Nov, Dec, Jan, ND, NDJ |
| Height of snow on Feb. 1st (cm) | Feb1HS | (observed value) | once/season |
| Height of snow on Mar. 1st (cm) | Mar1HS | (observed value) | once/season |

Strong temperature gradients (>10°C/m) through the snowpack are indicative of constructive metamorphism (Armstrong, 1985; McClung and Schaerer, 2006). Therefore, an average, minimum, and maximum temperature gradient were derived for each season. Each month or aggregate of months had a metric derived for average temperature gradient as well as a metric for the proportion of days with a negative temperature gradient greater than 10°C/m and a snow depth less than 1m. The latter metric was not included for February or March because the proportion was virtually zero in every

season, and deep persistent weak layers are not developed during these months. The proportion of days with a temperature gradient strong enough for constructive metamorphism was restricted to days with snow depths less than 1m due to attenuation of the temperature gradient from the top of the snowpack to the bottom (e.g., Gray and Male, 1981 p. 298; Armstrong and Brun, 2008 p. 39). The strength of the temperature gradient, which changes due to daily fluctuations in air temperature, generally decreases with distance from the snow surface (e.g., Bradley et al., 1977; Birkeland et al., 1998). Therefore, depth hoar layers are not commonly affected by temperature gradient metamorphism when the snowpack is sufficiently deep.

The calculated temperature gradient is not an exact measurement of the temperature gradient that the depth hoar layer is subject to. Two factors contribute to this fact. The depth of snow, as mentioned above, attenuates the air temperature, which results in a non-linear temperature gradient through the snowpack. In addition, longwave radiation loss can have a great effect on the snow surface temperature and is not commonly measured (Schweizer et al., 2003; Tremper, 2008). Therefore, the calculated temperature gradient, using air temperature, is an approximate estimate of when constructive metamorphism may be occurring at the base of the snowpack, and is included as an index of depth hoar formation. Metrics that summarize snow depth were derived, in part because a shallow snowpack can have constructive metamorphism when nighttime air temperatures are higher (i.e., weaker calculated temperature gradients) due to outgoing longwave radiation cooling the surface (Schweizer et al., 2003; Tremper, 2008).

The average and maximum snow depths were derived for each season and the individual months of November, December, and January (Table 12). Snow depth on February 1st and March 1st were the only snow depth metrics used for those months. These were chosen over average or maximum snow depth for those months because in a forecasting scenario the average and maximum are not known prior to avalanches that happen in those months, which are the subject of this research.

Total, maximum, and minimum new snowfall and SWE were derived for each season, and total new snowfall and SWE were derived for individual and aggregations of months. Totals of new snow and new SWE were used because the addition of load affects the development of buried persistent weak layers. Added weight can create more avalanches that remove or alter persistent weak layers early in the season, or compaction from overburden pressure can strengthen persistent crystals through pressure sintering (Brown et al., 2001).

Temperatures were summarized for each season by the extreme values (maximums and minimums) of the maximum and minimum temperature (Table 12). These were used in anticipation of a positive correlation between colder years and a weaker snowpack, or warmer years and deep wet slab activity on persistent weak layers.

### 3.4.4 Statistical Analysis Methods

The Wilcoxon rank-sum test (Ramsay and Schafer, 2002) was the univariate analysis performed to compare all seasonal metrics between seasons with deep slabs on persistent weak layers and seasons without deep slabs on persistent weak layers. It was also used to compare all observed and derived daily metrics between days with and days

without deep slabs on persistent weak layers. A p-value lower than 0.05 was considered strong evidence that a given metric's central tendency (i.e., mean) is different between seasons (or days) with and seasons (or days) without deep slabs on persistent weak layers.

Classification trees were grown and pruned through recursive partitioning using the rpart package version 4.1.1 (Therneau et al., 2013) in R statistical software version 3.0.0 (R Core Team). All trees were grown using the Gini index to create splits that reduce the probability of misclassification (Breiman et al., 1993). Over-fit (maximum sized) trees were grown using the software's default criteria to stop splitting when groups had less than 20 observations. These trees were used as exploratory tools to discuss the structure of meteorological data associated with deep slabs that fail on persistent weak layers late in the season. Over-fit trees have been applied for exploratory purposes as the structure of the maximum tree may reflect existing relationships between physical processes (Davis et al., 1999). However, over-fit trees are often over optimistic and not suited for prediction (Hendrikx et al., 2005; James et al., 2013). A large tree can contain splits that have poor predictive power on independent samples, and lead to higher true misclassification rates than a smaller "best" sized tree (Breiman et al., 1993). Therefore, a traditional cross-validation pruning approach was used to find the "best" sized tree, which would reflect more accurate measures of misclassification should the tree be used for prediction.

Classification trees have previously been implemented for avalanche forecasting purposes using 10-fold cross validation to find the "best" sized trees (e.g., Hendrikx et al., 2005; Baggi and Schweizer 2009; Peitzsch et al., 2012; Hendrikx et al., 2014). This

approach is described in detail in Breiman et al. (1993), and in regards to an avalanche dataset in Hendrikx et al. (2005). Previous work that has used this approach for avalanche datasets has produced trees with 71-86% overall correct classification (e.g., Baggi and Schweizer, 2009; Hendrikx et al., 2005) (Table 3). This method grows trees of each size from ten different random subsamples of the dataset. An average misclassification cost is calculated for each sized tree based on the prediction of observations from a random selection of test samples. This average misclassification cost is also referred to as the cross-validated (CV) cost (Hendrikx et al., 2005). The "best" sized tree, according to Breiman et al. (1993), is the least complex tree with a CV cost that does not differ significantly from the tree with the smallest CV cost. The one standard error rule was implemented to determine the least complex "best" sized tree (Breiman et al., 1993). This rule states that the "best" sized tree is the smallest sized tree whose CV costs do not differ by more than one times the standard error of the tree with the smallest CV cost (Breiman et al., 1993; Henrikx et al., 2005).

Daily and seasonal classification trees were created with a variety of explanatory variable datasets. Trees were created to predict three different response variables:

I. Trees that separate seasons with deep slabs on persistent weak layers and seasons without deep slabs on persistent weak layers. These trees were grown using three stratifications of the seasonal variables (Table 12), which each consisted of metrics from a different time period during the season. This allowed for trees to be fit for different forecasting situations (i.e., at certain times of the season prior to activity), as well as trees that summarize the characteristics of the entire season and provide general information about

deep slabs on persistent weak layers. The three subsets of metrics for fitting seasonal classification trees were:

1) Prior to January 1st (i.e., all seasonal metrics from November and December),
2) Prior to February 2nd (i.e., all seasonal metrics from November through January, and February 1st snow depth), and
3) All 45 seasonal metrics (Table 12).

A dataset that was restricted to the metrics found to have significant differences in the univariate analysis was not used as has been done previously (e.g., Judson and Erickson, 1973; Peitzsch et al., 2012). This was due in part to the various stratifications of response variables used, and in part to metrics potentially becoming significant after others are used in a tree.

II. Trees were also grown that split days with and days without deep slabs on persistent weak layers when a deep persistent weak layer was present and avalanche control was performed. Daily and seasonal metrics were used to make splits. Three different sets of metrics were used to grow trees that split days with and days without deep slabs on persistent weak layers:

1) A set of only daily metrics (tables 4 & 11),
2) A set of all daily and seasonal metrics (Table 12), and
3) A set of daily metrics that were found to have significantly different medians and distributions between days with and days without deep slabs on persistent weak layers, based on the Wilcoxon rank-sum test.

III. Trees were also fit with each of the previous three datasets to split days with deep dry slabs on persistent weak layers, deep wet avalanches on persistent weak layers, and no deep slabs on persistent weak layers.

Over-fit and cross-validated trees were fit for each dataset of seasons or days with deep slabs on persistent weak layers. Model performance of the cross-validated classification trees, and the over-fit trees, was described using measures explained by Wilks (1995) and Doswell et al. (1990), which have also been used previously by Schweizer et al. (2009) and Hendrikx et al. (2014). We use the unweighted average accuracy (RPC), the true skill score (TSS), the false alarm ratio (FAR), the probability of detection (POD), the probability of non-events (PON) and the probability of non-detection (FSR, i.e., false stable ratio). Ideal models have a low FAR and FSR, and a high POD and PON, which would lead to a high RPC and TSS (Hendrikx et al., 2014). These measures are defined in equations 1-6:

Unweighted average accuracy: $$RPC = 0.5\left(\frac{a}{a+c} + \frac{d}{b+d}\right) \quad (1)$$

True skill score: $$TSS = \frac{d}{b+d} - \frac{c}{a+c} \quad (2)$$

False alarm ratio: $$FAR = \frac{c}{c+d} \quad (3)$$

Probability of detection: $$POD = \frac{d}{b+d} \quad (4)$$

Probability of non-events: $$PON = \frac{a}{a+c} \quad (5)$$

Probability of non-detection: $$FSR = \frac{b}{b+d} \quad (6)$$

where the definitions for a-d are defined in a contingency Table (Table 13).

Table 13 Contingency table showing definitions for equations used to measure classification tree performance.

| | | Observed | |
|---|---|---|---|
| | | Non Avy | Avy |
| Predicted | No Avy | a = correct non-events | b = misses |
| | Avy | c = false alarms | d= hits |

These measures were slightly modified to assess model performance for the trees created to split deep slabs by type since there are three response variables in these trees. Correctly predicted non-event days were the population of correct non-events and false alarms were non-event days predicted as either type of event. All days that had either deep dry slabs or deep wet avalanches, on persistent weak layers, incorrectly classified were considered misses (Table 13). Days with either type of avalanche correctly predicted were considered hits (Table 13).

Random forests were created using the randomForest package in R (Liaw and Wiener, 2002) to iteratively grow 500 classification trees while successively withholding (with replacement) random variables and subsets of data. Based on the default methods in R, about one-third of the data is withheld from each tree, and the number of variables withheld from each tree is the square root of the number of variables in the dataset. One random forest was grown to predict days with deep slabs on persistent weak layers using only daily metrics (Table 4 &11), and one random forest was grown to predict seasons with deep slabs on persistent weak layers using only metrics prior to February 1$^{st}$ (Table 12). Model performance measures (eq. 1-6) for the random forest models were derived from the estimated classification rates for the data that was withheld from fitting of each tree, which is referred to as the "out-of-bag" sample (Breiman, 2001). Importance of each

variable was assessed using the mean decrease in accuracy (MDA). Each variable is individually removed from every tree in the forest, and then the MDA is calculated as the average decrease (across all trees) in error rate for classifying the out-of-bag data, when that variable is removed from a tree (Breiman, 2001).

# 4. RESULTS

The following section presents the findings of the case study followed by the results of the historical analysis of deep slabs on persistent weak layers. The case study provides a detailed account of an historical season that had both dry and wet deep slabs on a persistent weak layer. The season's meteorological history and the conditions that preceded days with deep slabs on a persistent weak layer provide a clear example of the association between the meteorological metrics that are used in the historical analysis and deep slabs on persistent weak layers.

## 4.1 Case Study: The 2011-2012 Season

During the 2011-2012 season, southwest Montana had many large avalanches and days of high avalanche danger (Chabot et al., 2012). In the Bridger Mountains the most prominent and active weak layer was a 30-55cm thick layer of depth hoar. This layer developed during the months of November, December, and January, and persisted through the season. Precipitation was below average between December and early February (55% of average), but each storm consistently led to explosive and human triggered avalanches on this layer. Higher snowfall rates and above average SWE in late February and March produced natural avalanche activity in the backcountry on the depth hoar layer. In late March 2012 a near isothermal snowpack, combined with a period of above freezing temperatures and heavy snowfall, produced an historic skier and explosive triggered full depth wet slab cycle on the depth hoar (Fig. 12). This event occurred on in-bounds terrain that was closed at the time, and had previously been heavily skied and

controlled with explosives. The following section is a review of meteorological and snowpack factors that were associated with avalanche activity on the depth hoar layer at Bridger Bowl during the 2011-12 season. Meteorological metrics derived from the Alpine weather station, and snowpack observations taken throughout the ski area, are used to discuss the factors associated with the timing of avalanche activity on the depth hoar layer.




Figure 12. Overview of the location (black outlines), and order (1-6), of large wet slab avalanches that failed on depth hoar on March 27th, 2012 at Bridger Bowl. Snowpack observation sites from the 2011-12 season are denoted with letters A-G (table 14).

4.1.1 Daily Weather and Snowpack History

Snow depth at the Alpine weather station was 30cm on November 30th after receiving 60mm of SWE (58% of monthly average) in the month of November. From December 1$^{st}$ to the 29$^{th}$, Alpine received just 26mm of additional SWE resulting in no significant increase in total snow depth (Fig. 13). Average temperature for both November and December was -3.3°C which is the historical average for November and 2.8°C above the historical average for December. Despite relatively warm temperatures, the shallow snowpack and cold, dry, and clear nights typical of this area were enough to

create a base of facets and depth hoar 20-30cm thick (Table 14) before the first significant snowfall began on December 30th (Fig. 13).

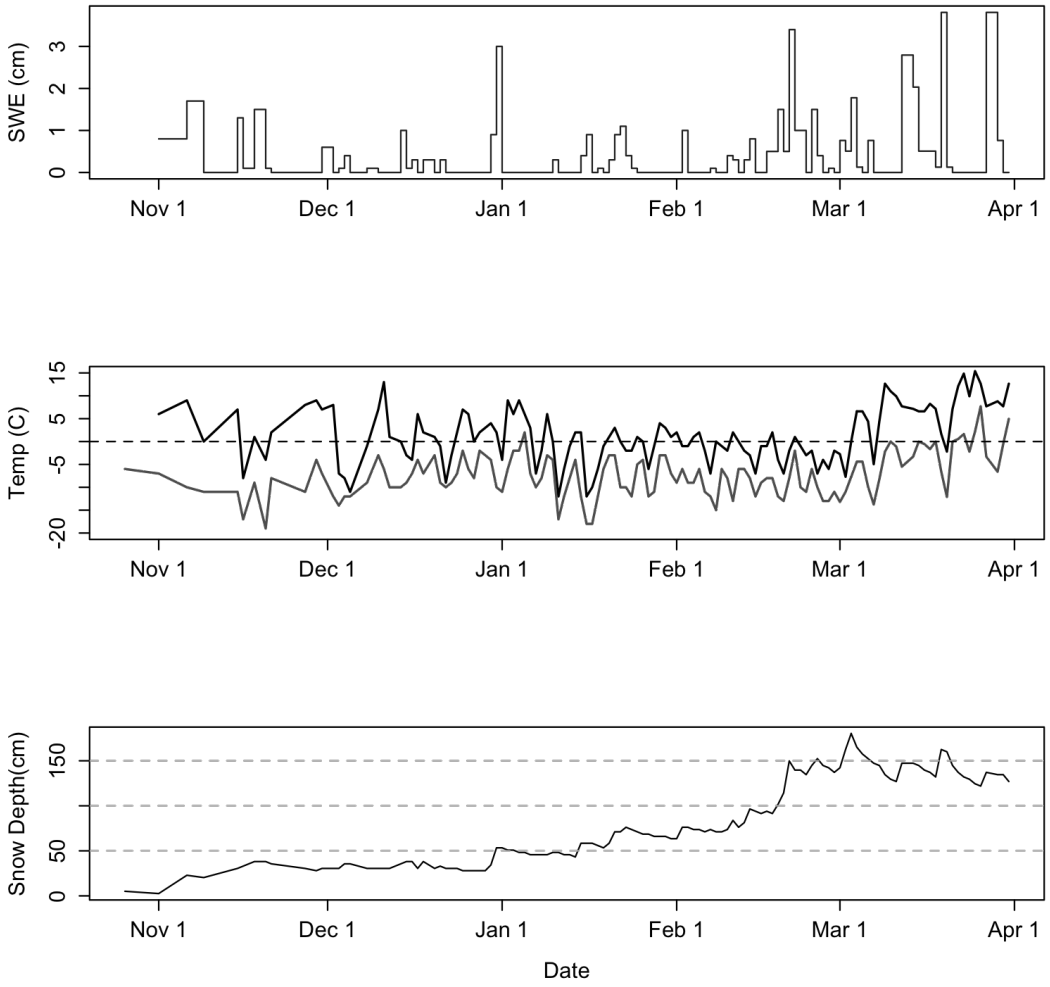



Figure 13. Meteorological observations from Bridger Bowl during the 2011-2012 season. (Top to bottom) Daily new snow water equivalent (SWE), daily minimum and maximum temperature, and snow depth.

Total SWE for December was doubled in a single storm (from 26mm to 65mm) between December 30th and January 1st, yet reached only 68% of the historical average for the month of December. The snow depth at Alpine was 58cm on January 31$^{st}$ (Fig. 13). Average temperature during January was -3.3°C (2.8°C above monthly average),

cold enough to allow for continued development of basal facets and depth hoar in the shallow snowpack (Table 14). Variability in and between the various areas of snowpack observation prevents any absolute quantification of this layer's development. However, observations of this layer's development show that it generally increased in thickness and grain size between early January and early February (Table 14).

Table 14 Snowpack observations during the 2011-2012 season at Bridger Bowl.

| Location (Fig.12) | Date | Snow depth (cm) | DH grain size (mm) | DH height (cm) | DH hardness | Slab thickness | Avg. slab hardness | ECT result |
|---|---|---|---|---|---|---|---|---|
| A | 15-Dec | 50 | 0.5 - 1.0 | 33 | 1 | 17 | 1.2 | N |
| A | 31-Dec | 75 | 2.0 - 3.0 | 27 | 1 | 48 | 1.3 | P |
| A | 3-Jan | 64 | 2.0 - 3.0 | 21 | 1 | 43 | 2.7 | P |
| A | 7-Jan | 75 | 2.0 - 3.0 | 30 | 1 | 45 | 1.7 | N |
| A | 18-Feb | 133 | 3.0 - 4.0 | 45 | 1 | 88 | 1.7 | P |
| B | 31-Dec | 76 | 1.0 - 2.0 | 33 | 2 | 43 | 1.3 | P |
| B | 3-Jan | 69 | 2.0 - 3.0 | 27 | 1 | 42 | 2.3 | N |
| C | 3-Jan | 84 | 1 | 45 | 1 | 39 | 1.5 | P |
| C | 7-Jan | 84 | 1.0 - 2.0 | 33 | 2 | 51 | 2.1 | P |
| C | 16-Jan | 95 | 3 | 45 | 0.85 | 50 | 1.4 | N |
| C | 22-Jan | 103 | 2 | 71 | 1 | 32 | 1.5 | P |
| D | 5-Jan | 84 | 2 | 41 | 1 | 43 | 1.8 | P |
| D | 10-Jan | 85 | 2.0 - 3.0 | 20 | 1 | 65 | 1.2 | N |
| E | 6-Jan | 74 | 2 | 34 | 1 | 40 | 1.2 | N |
| E | 16-Jan | 54 | 3.0-5.0 | 33 | 0.87 | 21 | 1.4 | N |
| E | 11-Feb | 137 | 3 | 45 | 1 | 92 | 1.6 | P |
| E | 1-Mar | 140 | 3.0-5.0 | 38 | 1.3 | 102 | 2.3 | P |
| F | 8-Mar | 180 | NA | 35 | 1 | 145 | NA | P |
| F | 11-Mar | 159 | 4.0 - 6.0 | 42 | 1 | 117 | 2.4 | P |
| F | 15-Mar | 195 | 2.0 - 3.0 | 50 | 1 | 145 | 2.3 | X |
| G | 20-Dec | 50 | 2.0 - 3.0 | 28 | 1 | 22 | 2.0 | N |
| G | 17-Mar | 210 | 3 | 55 | 1 | 155 | 2.1 | P |
| G | 24-Mar | 205 | 2 | 30 | 2 | 175 | 2.1 | X |

Precipitation rates increased in early February and temperatures were some of the lowest of the season (Fig. 13). Continued heavy precipitation over the last nine days of February brought the monthly amount of SWE to 134mm (168% of the monthly average). This trend continued between March 1st and 20th (Fig. 13). Snowpits at higher elevations show depths up to 210cm around this time with depth hoar still at the base of the snowpack (e.g., Location G, Table 14). While fractures were hard to initiate in extended column tests due to the great depth of the slab, they still produced sudden clean collapses on the depth hoar (Fig. 14).

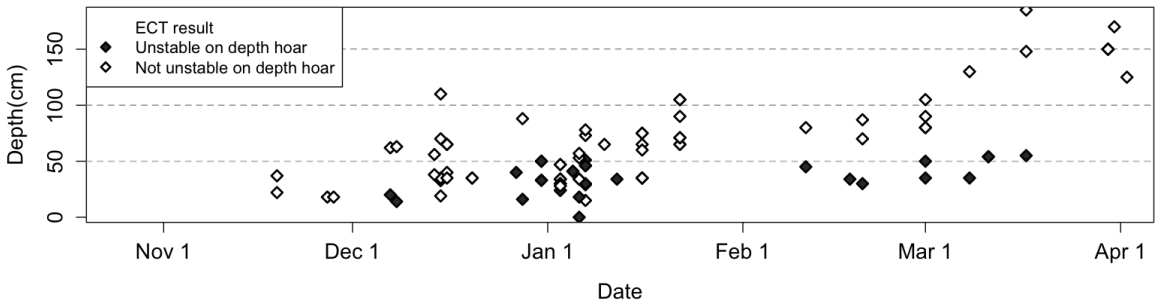



Figure 14. ECT results plotted by depth of failure layer (0 cm is the ground) during the 2011-2012 season. Results that were unstable (i.e., propagated during an ECT test) on the depth hoar layer are marked with dark points. Other test results (white points) include stable results on depth hoar as well as stable and unstable results on other layers.

Average temperatures in March were 1.7°C (4.5°C above monthly average). Minimum daily temperatures of 0°C or higher were recorded on the nights of March 9th, 14th and 17th. The first occurrence of multiple nights with above freezing minimum temperatures at the Alpine weather station began on the night of March 20$^{th}$ (Fig. 13). Between March 20th and 25th there was only one night with below freezing temperatures, and on the night of the 26th temperatures dropped below 0°C and 38mm of SWE fell overnight (Fig. 13).

The exact date when the snowpack became isothermal was not documented. However, a snowpit dug on March 17th had measured temperatures greater than -1.0°C in the middle of the snowpack, and a snowpit dug on March 24th, nearby a path that later produced a wet slab on the depth hoar layer on March 27th, had a temperature of -0.1°C measured at 100cm below the surface, and slightly colder temperatures in the top 20cm of the snowpack (snow profiles in Appendix B).

### 4.1.2 Avalanche Activity

Small storms in mid-December were enough to create small artificially triggered avalanches on the depth hoar layer (Fig. 15a). Enough snow fell on December 30th and 31st to create a dense slab that resulted in large artificially triggered slab avalanches on the depth hoar layer through January 6th despite no new snow after January 1st (Fig. 15b).

From then until February 24th, very few, if any, avalanches were observed on the depth hoar until a skier triggered an avalanche on Saddle Peak to the south of the ski area. Two days later, a natural avalanche ran on the depth hoar layer near this same area (Fig. 15c). Avalanches continued to occur naturally on the depth hoar in the backcountry, and a large deep slab cycle on the morning of March 20th marked the last dry slab activity observed in the backcountry on this layer (Fig. 3 & 15d).

Wet slabs were observed in the backcountry as early as March 23rd when a snowmobiler triggered a slide near Ross Peak, 6.5 km north of Bridger Bowl (Fig. 15e). Multiple wet slabs were recorded on March 26th including one spanning multiple paths directly north of the ski area boundary, and two on slopes near the slide recently triggered

by a snowmobiler (Fig. 15f). Numerous wet slabs were triggered in the new snow on March 27th, and six historic wet slabs were triggered on the depth hoar layer (Fig. 12 & 15g). These avalanches were relative size four and five, and reached historically observed maximum runout distances. One avalanche had an alpha angle of 19.5°, calculated using elevations and distance on Google Earth. Average alpha angles for maximum avalanche runout distances are typically between 20° and 28° (e.g., Mears, 1989; Delparte et al., 2008).

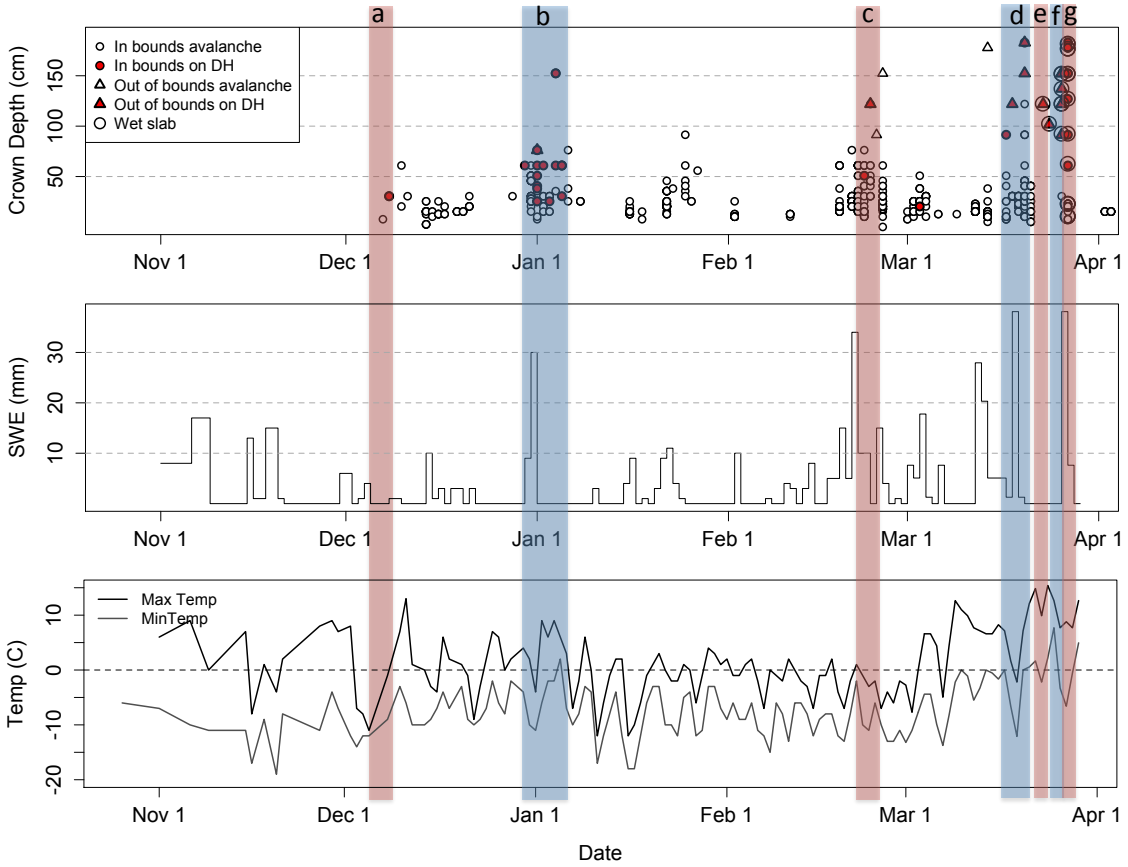



Figure 15. Daily SWE, temperature, and avalanche activity during the 2011-2012 season at Bridger Bowl. Avalanche activity is plotted (top panel) by depth and whether or not it slid on the depth hoar (red points), in bounds (circles) or out of bounds (triangles). Wet slabs are circled. Main events described in the text are shaded and labeled a-f.

4.1.3 Seasonal Meteorological Summary

Seasonal meteorological metrics that were believed to be associated with the formation of avalanches that fail on depth hoar were derived for all 44 seasons. A total of 45 metrics were of interest that summarize the meteorological conditions during the time that depth hoar forms and metamorphoses (Table 12), and may therefore be indicative of how reactive a depth hoar layer is later in the season.

The majority (80%) of values derived for seasonal meteorological metrics in the 2011-12 season were outside the interquartile range for all 44 seasons, and ten of the 45 metrics had an extreme value (minimum or maximum) in the 2011-12 season (Table 15). The metrics that had values greater than the interquartile range, or had their maximum value, in the 2011-12 season were measures of temperature, maximum temperature gradients, and counts of days with temperature gradients suitable for depth hoar formation. Metrics with values lower than the interquartile range, or having their minimum, in the 2011-12 season were mostly measures of snow depth, SWE, or temperature gradients (Table 15).

While the minimum temperature gradient during the 2011-12 season was the lowest for all seasons, the maximum temperature gradient was above the third quartile of all seasons. The average temperature gradient for November was a minimum value, and the average temperature gradients in December, January, and February were below the first quartile. The average temperature gradient in March was a maximum value. The proportions of days that had temperature gradients less than 1°C/10cm and snow depths less than 1m were above the third quartile in all months. November had a maximum proportion with all days meeting this criteria, and the proportion of days that met this

criteria in all three months combined (November, December, and January) was a maximum (Table 15).

Annual average height of snow, January height of snow, and height of snow on February 1st were all minimums in the 2011-2012 season. Heights of snow values in November and December, and on March 1st, were also below the first quartile. Total snowfall was within the interquartile range for November, but was below the interquartile range in December and January. Combined total snowfall in November, December, and January was also below the first quartile. Total SWE was within the interquartile range for November and December, and below the first quartile for January. Combined total SWE in November, December, and January was below the first quartile as well. All temperature metrics were above the third quartile with the maximum daily minimum temperature having a maximum value in the 2011-12 season (Table 15).

Table 15 Seasonal meteorological metrics for 2011-2012 compared to metrics for all 44 seasons. The -/+ indicates a value is below or above the first or third quartile, respectively. A * indicates a value from 2011-2012 is a minimum (-*) or maximum (+*) out of all seasons for that metric.

| Metric | Measured for 2011-12 | -/+ | All Seasons Min. | Median | Max. |
|---|---|---|---|---|---|
| Avg. height of snow (cm) | 79.2 | -* | 79.2 | 132.0 | 237.9 |
| Max. height of snow (cm) | 180.3 | - | 127.0 | 238.8 | 330.2 |
| Avg. 24 hr snow water equivalent (cm) | 0.4 | | 0.1 | 0.3 | 0.6 |
| Total 24 hr snow water equivalent (cm) | 49.3 | | 17.4 | 46.2 | 80.8 |
| Max. 24 hr snow water equivalent (cm) | 3.8 | | 1.5 | 3.8 | 8.8 |
| Avg. 24 hr snowfall (cm) | 4.1 | | 2.4 | 4.5 | 6.5 |
| Total 24 hr snowfall (cm) | 518.7 | - | 356.7 | 654.1 | 1009.2 |
| Max. 24 hr snowfall (cm) | 40.6 | | 25.4 | 43.5 | 109.2 |
| Min. temperature gradient (C/10cm) | -27.6 | -* | -27.6 | -3.8 | -1.6 |
| Avg. temperature gradient (C/10cm) | -1.7 | - | -1.8 | -0.8 | -0.5 |
| Max. temperature gradient (C/10cm) | 0.7 | + | -0.3 | 0.1 | 1.6 |
| Max. maximum temperature C | 15.0 | + | 4.0 | 13.0 | 19.0 |
| Min. maximum temperature C | -12.0 | + | -34.0 | -18.0 | -9.0 |
| Max. minimum temperature C | 8.0 | +* | -2.0 | 2.0 | 8.0 |
| Min. minimum temperature C | -19.0 | + | -42.0 | -28.0 | -15.0 |
| Avg. Nov. temperature gradient (C/10cm) | -5.7 | -* | -5.7 | -1.6 | -0.6 |
| Nov. Max. height of snow (cm) | 38.1 | - | 0.0 | 86.4 | 165.1 |
| Nov. Avg. height of snow (cm) | 28.2 | - | 0.0 | 49.0 | 119.9 |
| Nov. Total snow water equivalent (cm) | 6.1 | | 0.0 | 7.6 | 19.8 |
| Nov. Total 24 hr snowfall (cm) | 63.4 | | 0.0 | 91.6 | 255.6 |
| Days in Nov. with TG>10C/m & HS<1m | 1.0 | +* | 0.0 | 0.6 | 1.0 |
| Dec. Avg. temperature gradient (C/10cm) | -2.3 | - | -2.4 | -1.1 | -0.6 |
| Dec. Max. height of snow (cm) | 53.3 | - | 51.3 | 134.6 | 228.6 |
| Dec. Avg. height of snow (cm) | 32.7 | - | 31.2 | 97.7 | 170.8 |
| Dec. Total snow water equivalent (cm) | 6.5 | | 1.8 | 8.5 | 22.9 |
| Dec. Total 24 hr snowfall (cm) | 60.0 | - | 39.6 | 128.9 | 290.9 |
| Days in Dec. with TG>10C/m & HS<1m | 0.8 | + | 0.0 | 0.3 | 1.0 |
| Jan. Avg. temperature gradient (C/10cm) | -1.4 | - | -1.5 | -0.7 | -0.5 |
| Jan. Max. height of snow (cm) | 76.2 | -* | 76.2 | 177.8 | 271.8 |
| Jan. Avg. height of snow (cm) | 57.1 | -* | 57.1 | 137.0 | 243.6 |
| Jan. Total snow water equivalent (cm) | 4.5 | - | 0.0 | 8.9 | 21.5 |
| Jan. Total 24 hr snowfall (cm) | 57.9 | - | 50.8 | 125.1 | 250.8 |
| Days in Jan. with TG>10C/m & HS<1m | 0.6 | + | 0.0 | 0.0 | 0.7 |
| Feb. Avg. temperature gradient (C/10cm) | -1.0 | - | -1.1 | -0.6 | -0.3 |
| Feb. 1 height of snow (cm) | 63.5 | -* | 63.5 | 165.1 | 274.3 |
| Mar. Avg. temperature gradient (C/10cm) | -0.2 | +* | -0.8 | -0.3 | -0.2 |
| Mar. 1 height of snow (cm) | 142.2 | - | 96.5 | 189.3 | 304.8 |
| ND Avg. temperature gradient (C/10cm) | -3.3 | - | -3.4 | -1.3 | -0.7 |
| ND Total snow water equivalent (cm) | 12.6 | | 5.5 | 17.7 | 33.4 |
| ND Total 24 hr snowfall (cm) | 123.4 | - | 91.4 | 236.6 | 439.4 |
| Days in Nov. & Dec. with TG>10C/m & HS<1m | 0.9 | + | 0.0 | 0.4 | 1.0 |
| NDJ Avg. temperature gradient (C/10cm) | -2.4 | - | -2.6 | -1.1 | -0.6 |
| NDJ Total snow water equivalent (cm) | 17.1 | - | 10.3 | 25.8 | 53.0 |
| NDJ Total 24 hr snowfall (cm) | 181.3 | - | 158.3 | 368.3 | 593.7 |
| Days in Nov. Dec. & Jan. with TG>10C/m & HS<1m | 0.8 | +* | 0.0 | 0.3 | 0.8 |

## 4.2 Historical Analysis: 1968/69-2012/13

### 4.2.1 Univariate Analysis Results

The Wilcoxon rank-sum test indicates that 25 of the 45 seasonal metrics have significantly different central tendencies between seasons with deep slabs on persistent weak layers and seasons without deep slabs on persistent weak layers (Table 16). Summaries of temperature over the entire season, metrics in the month of March, and all but one metric in November (count of days in November with a temperature gradient >10°C/m and snow depth <1m) did not show significant differences at the $p<0.05$ level. All snow depth, new snow, and SWE summaries for the entire season, except for maximum new snow, did show significant differences at $p<0.05$. Average temperature gradient for the entire season, and in both January and February, was also significantly different between seasons at this level. Metrics that were significantly different between seasons for the month of December included maximum snow depth, total SWE, and total new snow. Average snow depth, total SWE, and total new snow in January were all significantly different between seasons with and without deep slabs on persistent weak layers. All seasonal metrics that summarized conditions over November and December, or over November, December and January were significantly different (Table 16).

Table 16 Summary statistics for all seasonal metrics and results from the Wilcoxon rank-sum test (p-value), used to test for a difference in central tendency between seasons with (DS) and without (No-DS) deep slabs on persistent weak layers after February 1st.

| Metric | Min. | | Median | | Max. | | p-value | p<.05= |
|---|---|---|---|---|---|---|---|---|
| n; DS=24 No-DS=20 | DS | No-DS | DS | No-DS | DS | No-DS | two-sided | Y |
| avgHS | 79.2 | 81.6 | 121.9 | 155.0 | 176.1 | 237.9 | 0.0071 | Y |
| maxHS | 162.6 | 127.0 | 221.0 | 257.8 | 299.7 | 330.2 | 0.0391 | Y |
| avgSWE | 0.1 | 0.2 | 0.3 | 0.4 | 0.4 | 0.6 | 0.0027 | Y |
| sumSWE | 17.4 | 24.6 | 43.0 | 54.6 | 61.5 | 80.8 | 0.0077 | Y |
| maxSWE | 1.5 | 2.0 | 3.3 | 4.2 | 6.1 | 8.8 | 0.0149 | Y |
| avgHN | 2.7 | 2.4 | 4.1 | 5.1 | 5.6 | 6.5 | 0.0007 | Y |
| sumHN | 397.6 | 356.7 | 569.1 | 722.6 | 849.9 | 1009.2 | 0.0082 | Y |
| maxHN | 25.4 | 26.2 | 41.9 | 47.0 | 106.7 | 109.2 | 0.3100 | |
| minTG | -27.6 | -13.7 | -4.6 | -3.2 | -2.3 | -1.6 | 0.0518 | |
| avgTG | -1.8 | -1.8 | -1.0 | -0.7 | -0.6 | -0.5 | 0.0048 | Y |
| maxTG | -0.1 | -0.3 | 0.2 | 0.1 | 0.8 | 1.6 | 0.6452 | |
| mxTmx | 4.0 | 7.0 | 12.5 | 13.6 | 19.0 | 19.0 | 0.3309 | |
| mnTmx | -24.0 | -34.0 | -18.0 | -19.0 | -12.0 | -9.0 | 0.1895 | |
| mxTmn | -2.0 | -2.0 | 2.0 | 2.0 | 8.0 | 7.0 | 0.3855 | |
| mnTmn | -33.0 | -42.0 | -28.0 | -29.0 | -19.0 | -15.0 | 0.7142 | |
| avgNovtg | -5.7 | -3.9 | -1.8 | -1.1 | -0.9 | -0.6 | 0.0522 | |
| maxHSnov | 0.0 | 38.1 | 72.8 | 90.2 | 130.3 | 165.1 | 0.1254 | |
| avgHSnov | 0.0 | 16.2 | 40.8 | 54.0 | 75.3 | 119.9 | 0.1355 | |
| sumSWEnov | 0.0 | 2.5 | 6.4 | 8.5 | 17.2 | 19.8 | 0.1437 | |
| sumHNnov | 0.0 | 34.3 | 75.4 | 103.0 | 182.9 | 255.6 | 0.1791 | |
| Novtgcnt | 0.0 | 0.0 | 0.6 | 0.5 | 1.0 | 1.0 | 0.0177 | Y |
| avgDectg | -2.4 | -2.0 | -1.2 | -1.0 | -0.6 | -0.6 | 0.1069 | |
| maxHSdec | 51.3 | 89.8 | 109.3 | 139.9 | 203.2 | 228.6 | 0.0463 | Y |
| avgHSdec | 31.2 | 65.7 | 82.1 | 106.8 | 146.4 | 170.8 | 0.0790 | |
| sumSWEdec | 1.8 | 5.7 | 6.0 | 10.6 | 22.9 | 21.7 | 0.0080 | Y |
| sumHNdec | 39.6 | 84.4 | 89.5 | 139.0 | 290.9 | 259.6 | 0.0209 | Y |
| Dectgcnt | 0.0 | 0.0 | 0.5 | 0.2 | 1.0 | 0.8 | 0.0517 | |
| avgJantg | -1.5 | -1.4 | -0.8 | -0.6 | -0.5 | -0.5 | 0.0233 | Y |
| maxHSjan | 76.2 | 111.8 | 150.4 | 184.2 | 251.5 | 271.8 | 0.0545 | |
| avgHSjan | 57.1 | 96.4 | 132.9 | 154.9 | 194.2 | 243.6 | 0.0383 | Y |
| sumSWEjan | 0.0 | 0.0 | 7.1 | 11.5 | 19.6 | 21.5 | 0.0236 | Y |
| sumHNjan | 50.8 | 66.3 | 110.6 | 149.3 | 219.7 | 250.8 | 0.0229 | Y |
| Jantgcnt | 0.0 | 0.0 | 0.0 | 0.0 | 0.7 | 0.3 | 0.0598 | |
| avgFebtg | -1.0 | -1.1 | -0.6 | -0.5 | -0.4 | -0.3 | 0.0063 | Y |
| Feb1HS | 63.5 | 99.1 | 147.3 | 172.7 | 261.6 | 274.3 | 0.0606 | |
| avgMartg | -0.7 | -0.8 | -0.3 | -0.4 | -0.2 | -0.2 | 0.9062 | |
| Mar1HS | 119.4 | 96.5 | 179.1 | 208.3 | 269.2 | 304.8 | 0.0770 | |
| NDavgtg | -3.4 | -2.9 | -1.5 | -1.2 | -0.8 | -0.7 | 0.0110 | Y |
| NDsumSWE | 5.5 | 8.2 | 13.1 | 19.2 | 28.3 | 33.4 | 0.0284 | Y |
| NDsumHN | 91.4 | 123.3 | 205.0 | 252.0 | 407.6 | 439.4 | 0.0431 | Y |
| NDtgcnt | 0.2 | 0.0 | 0.6 | 0.3 | 1.0 | 0.9 | 0.0011 | Y |
| NDJavgtg | -2.6 | -2.4 | -1.3 | -1.0 | -0.7 | -0.6 | 0.0071 | Y |
| NDJsumSWE | 10.3 | 15.9 | 23.1 | 31.4 | 47.9 | 53.0 | 0.0021 | Y |
| NDJsumHN | 158.3 | 228.5 | 318.2 | 387.1 | 590.0 | 593.7 | 0.0095 | Y |
| NDJtgcnt | 0.1 | 0.0 | 0.4 | 0.2 | 0.8 | 0.7 | 0.0006 | Y |

There were 27 daily metrics that showed significantly different central tendencies between days with deep slabs on persistent weak layers and days without deep slabs on persistent weak layers, when a deep persistent weak layer was present. These included rainfall over the previous 24-hours and all temperature metrics except for the change in maximum or minimum temperature over two days (Table 17). The daily metrics that measure new load (i.e., New SWE or snow over 1, 2, 3, 5, or 7 days), or snow depth change, were not significantly different at the $p<0.05$ level. Although, some metrics show p-values just slightly over $p=0.05$, and may become significant after accounting for other metrics in a multivariate analysis (Table 17).

### 4.2.2 Classification Trees for Seasons with Deep Slabs on Persistent Weak Layers

The largest classification tree grown to split seasons with deep slabs on persistent weak layers and seasons without deep slabs on persistent weak layers using all metrics prior to January 1st had three splits. These splits used the following metrics: sum of new snow in December, average temperature gradient in December, and the proportion of days in November with a temperature gradient greater than 10°C/m and snow depth less than 1m (Fig. 16). This tree achieved a POD of 0.96 with a FSR of 0.04, a FAR of 0.18, and a RPC of 0.85 (Table 18). The cross-validated tree using metrics before January 1st has one split that uses the metric for sum of new snow in December (Fig. 17). The performance measures for this tree are lower than the over-fit tree, with a POD of 0.58, a FSR of 0.42, and a RPC of 0.77. The FAR remains low at 0.07 (Table 18).

Table 17 Summary statistics for all daily metrics, and results of Wilcoxon rank-sum test (p-value), which was used to test for a difference in central tnedency between days with (DS) and days without (No-DS) deep slabs on a persistent weak layer.

| Metric | Min. | | Median | | Max. | | P-value | p < .05=Y |
|---|---|---|---|---|---|---|---|---|
| n; DS=76 | | | | | | | two-sided | |
| No-DS=76 | DS | No-DS | DS | No-DS | DS | No-DS | | |
| HS | 101.6 | 83.8 | 172.7 | 172.7 | 246.4 | 299.7 | 0.9187 | |
| HS2 | -12.7 | -15.3 | -1 | 2.5 | 54 | 53.4 | 0.0540 | |
| HS3 | -20.3 | -17.8 | 0 | 5.1 | 87 | 63.5 | 0.0553 | |
| HS5 | -30.5 | -33 | 6.35 | 10.2 | 97 | 78.8 | 0.1499 | |
| HS7 | -38.1 | -45.7 | 12.7 | 10.2 | 94 | 73.7 | 0.5510 | |
| HN | 0 | 0 | 2.5 | 5.1 | 55.9 | 40.6 | 0.0674 | |
| HN2 | 0 | 0 | 12.05 | 14 | 88.9 | 68.6 | 0.0823 | |
| HN3 | 0 | 0 | 15.2 | 20.3 | 90.2 | 75 | 0.0541 | |
| HN5 | 0 | 0 | 28.75 | 31.1 | 105.4 | 109.2 | 0.7368 | |
| HN7 | 0 | 0 | 40.6 | 41.7 | 154.9 | 143.5 | 0.6102 | |
| SWE | 0 | 0 | 0.1 | 0.3 | 3.8 | 4.6 | 0.1588 | |
| SWE2 | 0 | 0 | 0.75 | 0.8 | 5.6 | 7.8 | 0.3322 | |
| SWE3 | 0 | 0 | 1.2 | 1.3 | 5.9 | 7.9 | 0.3908 | |
| SWE5 | 0 | 0 | 2.3 | 2 | 8.1 | 8 | 0.6731 | |
| SWE7 | 0 | 0 | 3 | 2.9 | 8.9 | 8 | 0.6526 | |
| Rn | 0 | 0 | 0 | 0 | 2.5 | 0 | 0.0184 | Y |
| Tmin | -23 | -27 | -7 | -10 | 8 | 1 | 0.0000 | Y |
| Tmin2 | -29 | -30 | -9 | -12 | 2 | 0 | 0.0000 | Y |
| Tmin3 | -29 | -30 | -9 | -12 | 2 | -3 | 0.0000 | Y |
| Tmin5 | -30 | -30 | -11 | -13 | 0 | -6 | 0.0004 | Y |
| Tmin7 | -30 | -33 | -12 | -14 | -3 | -6 | 0.0026 | Y |
| Tminav2 | -26 | -26 | -6.5 | -10 | 5 | 0.5 | 0.0000 | Y |
| Tminav3 | -27 | -25 | -7.3 | -9.3 | 2.7 | -0.7 | 0.0000 | Y |
| Tminav5 | -27.6 | -23.8 | -7.6 | -10.15 | 2.2 | -2 | 0.0001 | Y |
| Tminav7 | -27.6 | -19.1 | -7.4 | -9.6 | 1.1 | -2.9 | 0.0002 | Y |
| TminDif2 | -16 | -14 | 0 | -1 | 11 | 13 | 0.8062 | |
| Tmax | -14 | -22 | 2 | -2 | 15 | 13 | 0.0000 | Y |
| Tmax2 | -14 | -21 | 3 | 1 | 16 | 13 | 0.0009 | Y |
| Tmax3 | -14 | -10 | 3 | 1 | 16 | 13 | 0.0003 | Y |
| Tmax5 | -14 | -7 | 4 | 3 | 16 | 13 | 0.0034 | Y |
| Tmax7 | -14 | -6 | 5 | 3 | 16 | 13 | 0.0067 | Y |
| Tmaxav2 | -15 | -21.5 | 1.5 | -2 | 15 | 9 | 0.0001 | Y |
| Tmaxav3 | -15.7 | -15.3 | 1.7 | -2.3 | 14.7 | 9.7 | 0.0000 | Y |
| Tmaxav5 | -16.8 | -12.6 | 1 | -1.9 | 13 | 9.2 | 0.0001 | Y |
| Tmaxav7 | -17.3 | -9.4 | 1.1 | -1.1 | 11.4 | 6.4 | 0.0004 | Y |
| TmaxDif2 | -12 | -16 | 0 | 0 | 10 | 12 | 0.9193 | |
| maxAbovez | 0 | 0 | 2 | 0 | 15 | 13 | 0.0005 | Y |
| maxAbovez2 | 0 | 0 | 4 | 1 | 30 | 18 | 0.0010 | Y |
| maxAbovez3 | 0 | 0 | 6 | 1 | 44 | 29 | 0.0003 | Y |
| maxAbovez5 | 0 | 0 | 9.5 | 5 | 65 | 46 | 0.0017 | Y |
| minAbovez | 0 | 0 | 0 | 0 | 8 | 1 | 0.0108 | Y |
| minAbovez2 | 0 | 0 | 0 | 0 | 10 | 1 | 0.0002 | Y |
| minAbovez3 | 0 | 0 | 0 | 0 | 10 | 1 | 0.0005 | Y |
| minAbovez5 | 0 | 0 | 0 | 0 | 13 | 1 | 0.0001 | Y |

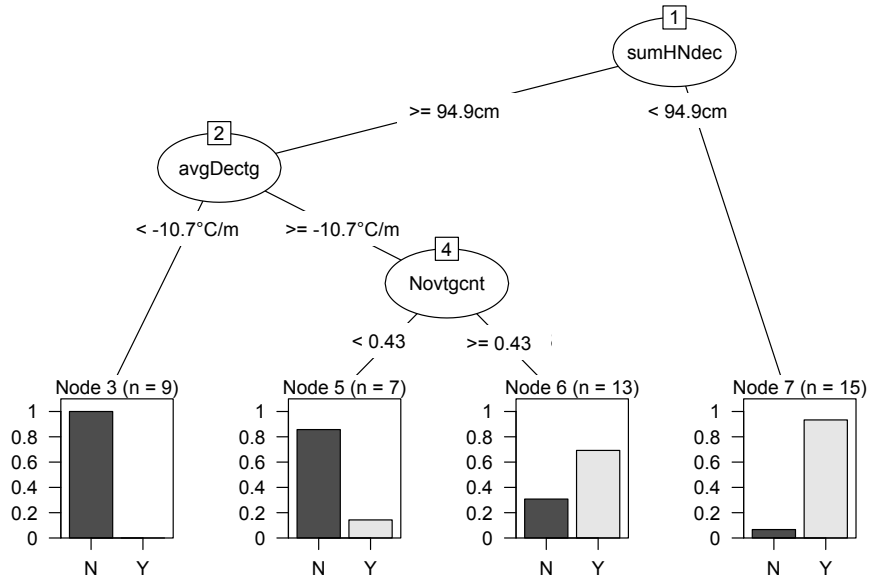



Figure 16. Over-fit classification tree created from seasonal metrics prior to January 1st to split seasons with (Y) and seasons without (N) deep slabs on persistent weak layers. (Tree graphics created using the partykit package in R statistical software (Hothorn and Zeileis, 2013)).

Table 18 Contingency table values (a-d) and performance measures for classification trees grown to split seasons with and seasons without deep slabs on persistent weak layers. Definitions for a,b,c, and d are given in table 13. Maximum (over-fit) and cross-validated tree results are both shown.

| Metrics used to grow tree | a | b | c | d | RPC | TSS | FAR | POD | PON | FSR |
|---|---|---|---|---|---|---|---|---|---|---|
| Over-fit trees | | | | | | | | | | |
| Pre Jan. 1st | 15 | 1 | 5 | 23 | 0.85 | 0.71 | 0.18 | 0.96 | 0.75 | 0.04 |
| Pre Feb. 2nd | 19 | 5 | 1 | 19 | 0.87 | 0.74 | 0.05 | 0.79 | 0.95 | 0.21 |
| All Metrics | 18 | 5 | 2 | 19 | 0.85 | 0.69 | 0.10 | 0.79 | 0.90 | 0.21 |
| Cross-validated trees | | | | | | | | | | |
| Pre Jan. 1st | 19 | 10 | 1 | 14 | 0.77 | 0.53 | 0.07 | 0.58 | 0.95 | 0.42 |
| Pre Feb. 2nd | 11 | 1 | 9 | 23 | 0.75 | 0.51 | 0.28 | 0.96 | 0.55 | 0.04 |
| All Metrics | 11 | 1 | 9 | 23 | 0.75 | 0.51 | 0.28 | 0.96 | 0.55 | 0.04 |

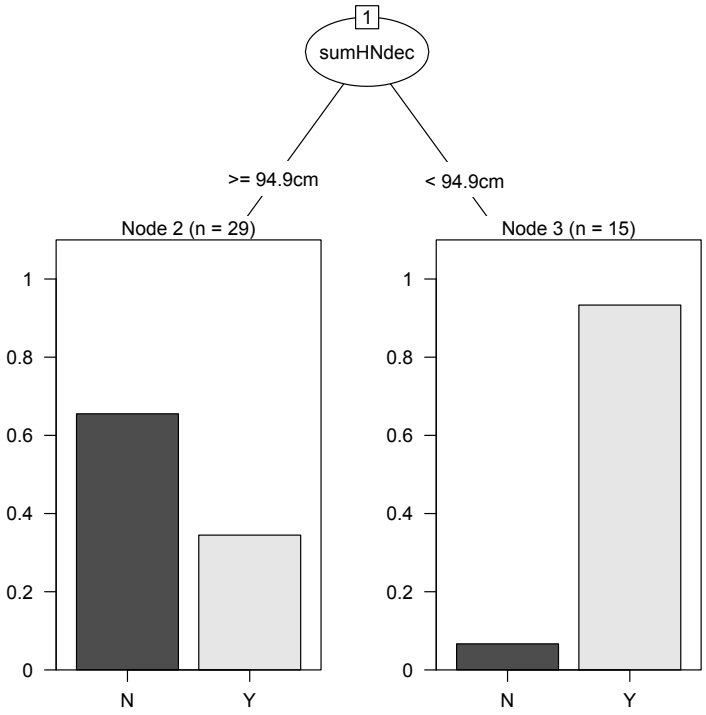



Figure 17. Cross-validated classification tree created from seasonal metrics prior to January 1st to split seasons with (Y) and seasons without (N) deep slabs on persistent weak layers.

The maximum tree using all metrics prior to February 2nd has a slightly higher RPC at 0.87 and a lower FAR at 0.05, but the POD drops to 0.79 with a FSR of 0.21 (Table 18). This tree has three splits that use the metrics: sum of SWE in November, December and January, sum of SWE in December, and maximum snow depth in November (Fig. 18). The cross-validated tree using metrics prior to February 2nd has one split using the metric for sum of SWE in November, December, and January (Fig. 19). This tree obtains a POD of 0.96, FSR of 0.04, FAR of 0.28, and RPC of 0.75 (Table 18).

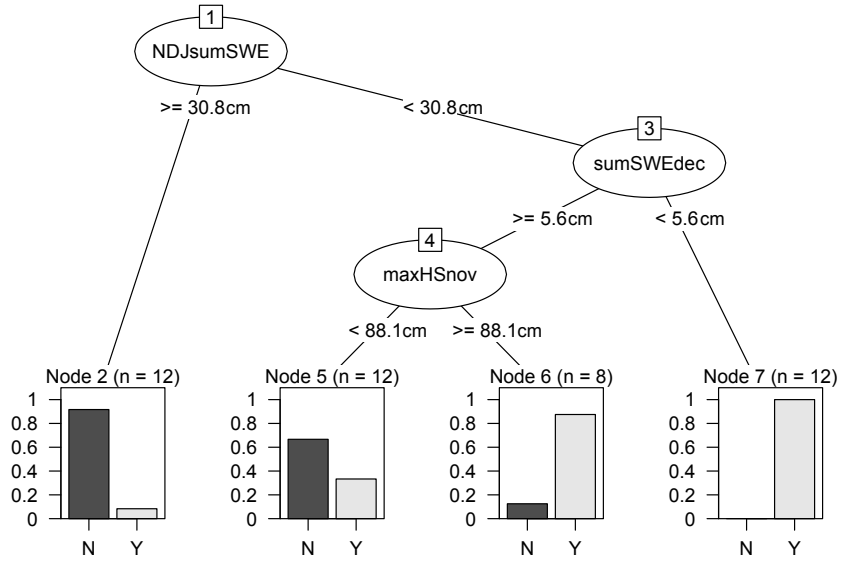



Figure 18. Over-fit classification tree using seasonal metrics prior to February 2nd to split seasons with (Y) and seasons without (N) deep slabs on persistent weak layers.

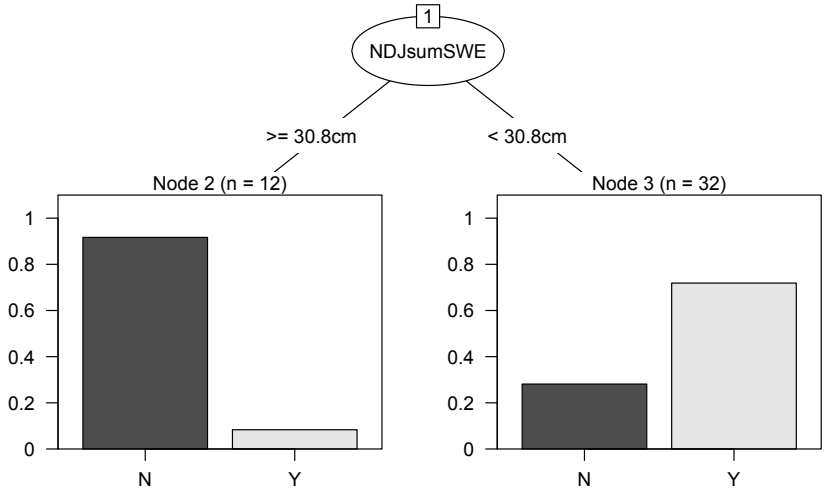



Figure 19. Cross-validated classification tree using all metrics prior to February 2nd to split seasons with (Y) and seasons without (N) deep slabs on persistent weak layers. The same tree is obtained when using metrics from the entire season.

The maximum tree using all seasonal metrics has two splits, which use the metrics: sum of SWE in November, December and January and the average temperature gradient in February (Fig. 20). This tree had an RPC of 0.85, POD of 0.79, FSR of 0.21,

and FAR of 0.10 (Table 18). The cross-validated trees using metrics from the entire season was the same as the cross-validated tree using metrics prior to February 2nd (Fig. 19).

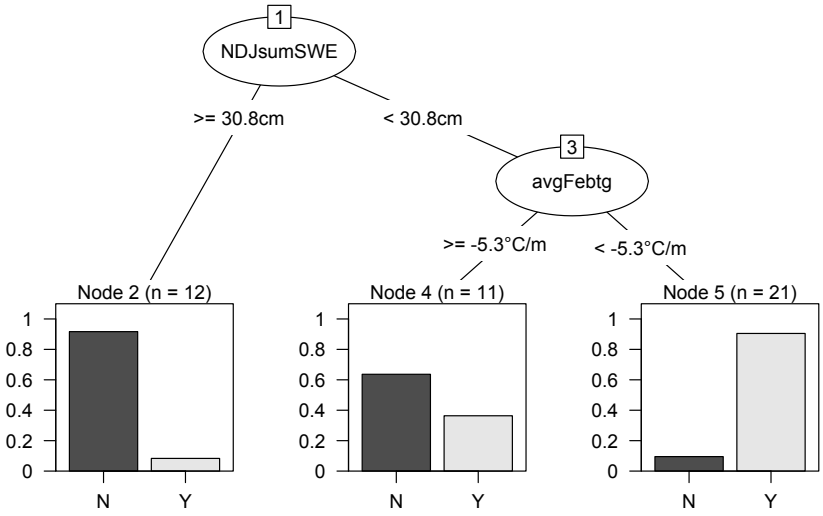



Figure 20. Over-fit classification tree using all seasonal metrics to split seasons with (Y) and seasons without (N) deep slabs on persistent weak layers.

### 4.2.3 Classification Trees for Days with Deep Slabs on Persistent Weak Layers

The over-fit classification tree that split days with and days without deep slabs on persistent weak layers using only daily metrics has a primary split based on minimum temperature over the past 24-hours (Fig. 21). The second node splits the group of observations that have colder minimum 24-hour temperatures, based on the total SWE over the previous five days. This creates a group of mostly avalanche days when there is more than 4.5cm of SWE over five days. The group of days with warmer minimum 24-hour temperatures were split by three-day average maximum temperature (Fig. 21). One group created from this split consists of avalanche days, when the three-day average maximum temperature is greater than 7.2°C. For days with a lower three-day average

maximum temperature a split is made based on the change in snow depth over the previous seven days. Days with more than 10cm of snow settlement over the previous seven days had zero deep slabs on persistent weak layers. Days with less settlement, or more load, are split up based on the height of snow, with more deep slabs on days that have greater than 164cm snow depth. Days with a lower snow depth are split up by the change in snow depth over the previous three days. The days that had an increase in snow depth less than 9cm had more days with deep slabs on persistent weak layers than days that had an increase in snow depth of 9cm or more (Fig. 21).

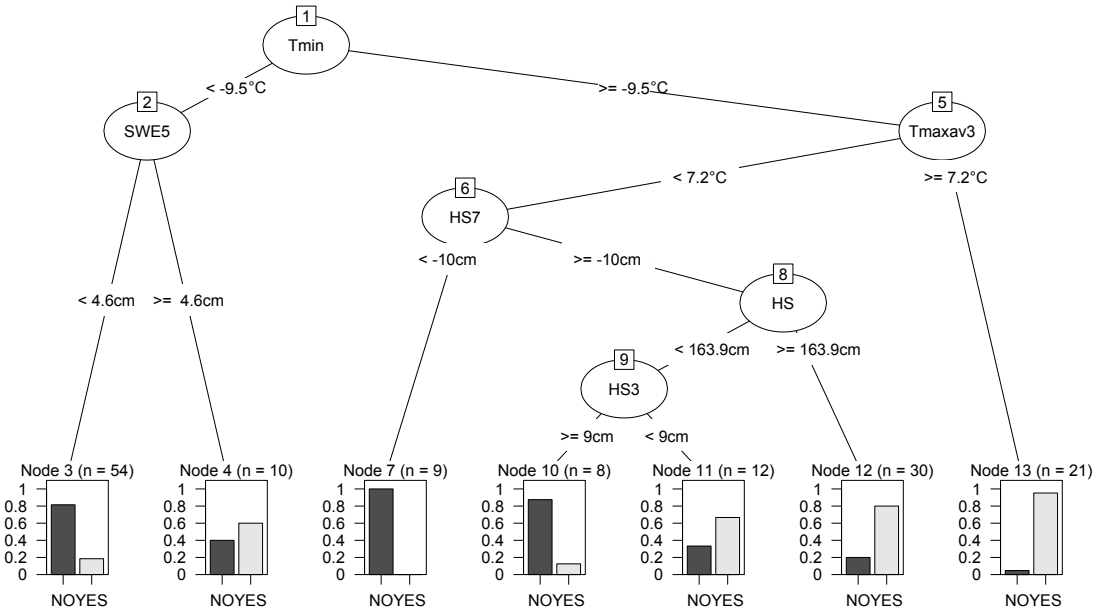



Figure 21. Over-fit classification tree grown using only daily meteorological metrics to split days with (YES) and days without (NO) deep slabs on persistent weak layers.

The cross-validated classification tree grown to split days with deep slabs on persistent weak layers from days without deep slabs on persistent weak layers using only daily metrics had three splits using the metrics: minimum 24-hour temperature, average maximum temperature over 72-hours, and snow depth change over seven days (Fig. 22).

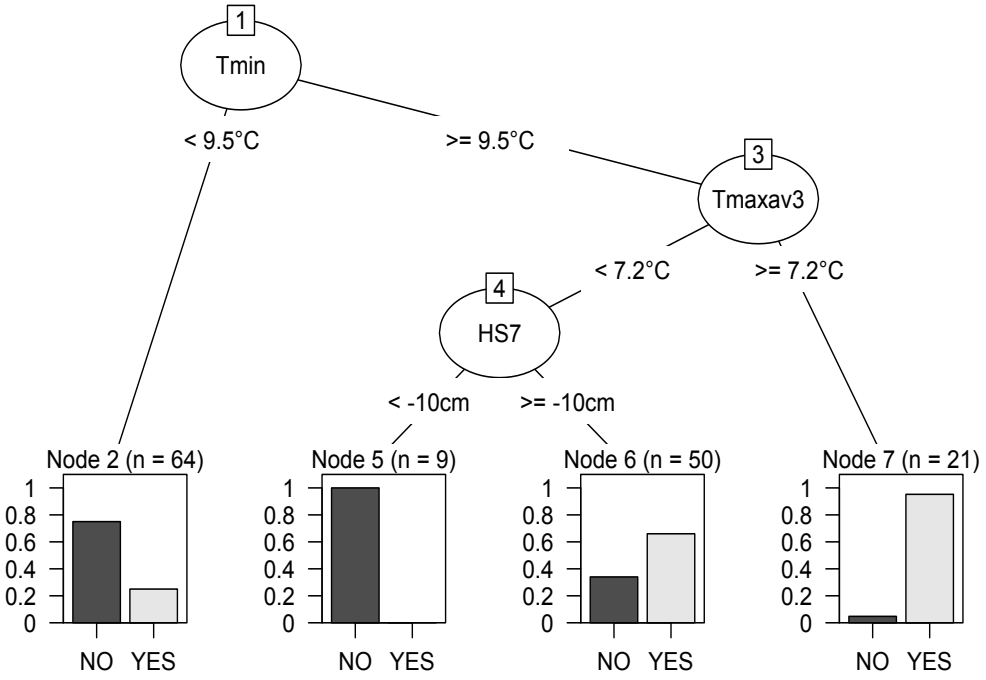



Figure 22. Cross-validated classification tree using daily metrics to split days with (YES) and days without (NO) deep slabs on persistent weak layers. The same tree is obtained when all daily and seasonal metrics are used.

The same tree was obtained when using daily metrics and all daily and seasonal metrics. This tree had a POD of 0.79, FSR of 0.21, FAR of 0.24 and RPC of 0.77 (Table 19). The tree using only daily metrics has slightly different performance measures due to missing daily values. Over-fit trees using all three sets of metrics (Appendix C) had more optimistic FARs as low as 0.14 and PODs up to 0.84 (Table 19). The cross-validated tree that used only metrics that were significantly different based on the Wilcoxon rank-sum test (Table 17) had one split that used 24-hour minimum temperature (Fig. 23). This tree had a POD of 0.77, FAR of 0.34, and RPC of 0.70 (Table 19).

Table 19 Contingency table values (a-d) and performance measures for classification trees grown to split days with from days without deep slabs on persistent weak layers. Definitions for a,b,c, and d are given in table 13. Maximum (over-fit) and cross-validated tree results are both shown.

| Metrics used to grow tree | a | b | c | d | RPC | TSS | FAR | POD | PON | FSR |
|---|---|---|---|---|---|---|---|---|---|---|
| Over-fit trees | | | | | | | | | | |
| All Daily Metrics | 60 | 11 | 15 | 58 | 0.82 | 0.64 | 0.21 | 0.84 | 0.80 | 0.16 |
| All Metrics | 66 | 14 | 10 | 62 | 0.84 | 0.68 | 0.14 | 0.82 | 0.87 | 0.18 |
| Only significant | 58 | 12 | 17 | 57 | 0.80 | 0.60 | 0.23 | 0.83 | 0.77 | 0.17 |
| Cross-validated trees | | | | | | | | | | |
| All Daily Metrics | 57 | 16 | 18 | 53 | 0.76 | 0.53 | 0.25 | 0.77 | 0.76 | 0.23 |
| All Metrics | 57 | 16 | 19 | 60 | 0.77 | 0.54 | 0.24 | 0.79 | 0.75 | 0.21 |
| Only Significant | 48 | 16 | 27 | 53 | 0.70 | 0.41 | 0.34 | 0.77 | 0.64 | 0.23 |

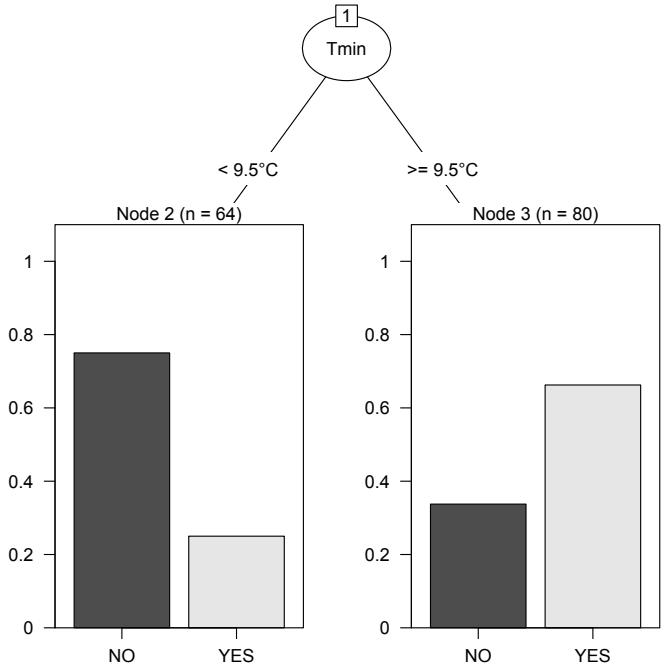



Figure 23. Cross-validated classification tree grown to split days with (YES) and days without (NO) deep slabs on persistent weak layers using only metrics with significantly different distributions between days with and days without deep slabs on persistent weak layers.

The cross-validated tree grown to split days with deep dry slabs on persistent weak layers, days with deep wet slabs on persistent weak layers, and days without deep slabs on persistent weak layers used the same splits as the tree that split only days with and days without deep slabs on persistent weak layers, but in a different order. This tree used average 72-hour maximum temperature instead of minimum 24-hour temperature as the first split, followed by 24-hour minimum temperature and change in snow depth over seven days (Fig. 24). Days with a three day average maximum temperature greater than 7.5°C had deep wet slabs. Days with a lower three day average maximum temperature had less deep slabs when the minimum 24-hour temperature was less than -9.5°C. When the minimum 24-hour temperature was greater, days with less settlement (more load) had deep dry slabs and days with less loading are mostly days without deep slabs (Fig. 24). Again, the same tree was obtained regardless of using only daily metrics or all daily and seasonal metrics. The cross-validated tree that used only variables that were significantly different between days with and days without deep slabs in the univariate analysis had only one split using the metric for average 72-hour maximum temperature (Fig. 25).

Model performance measures for the trees that split deep slabs on persistent weak layers by type were lower overall than measures for cross-validated trees that split days with and days without deep slabs on persistent weak layers (Table 19 & 20). The RPC and FAR are similar to the other cross-validated trees, at 0.73 and 0.25, respectively. However, the POD and FSR are around 0.1 points worse (lower POD and higher FSR) than the other trees at 0.68 and 0.32, respectively (Table 20). Performance measures for the tree that used only significant metrics to split days with deep slabs on persistent weak layers by type were poor overall with a POD of 0.26 and an FSR of 0.74 (Table 20).

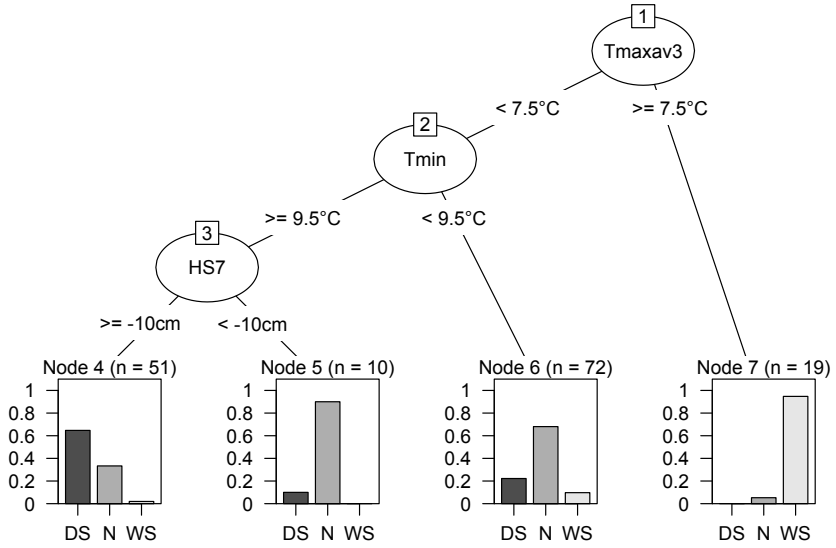



Figure 24. Cross-validated classification tree obtained using the datasets: all daily metrics and daily and seasonal metrics. This tree splits days with deep wet avalanches on persistent weak layers (WS), days with deep dry slabs on persistent weak layers (DS), and days without deep slabs on persistent weak layers (N).

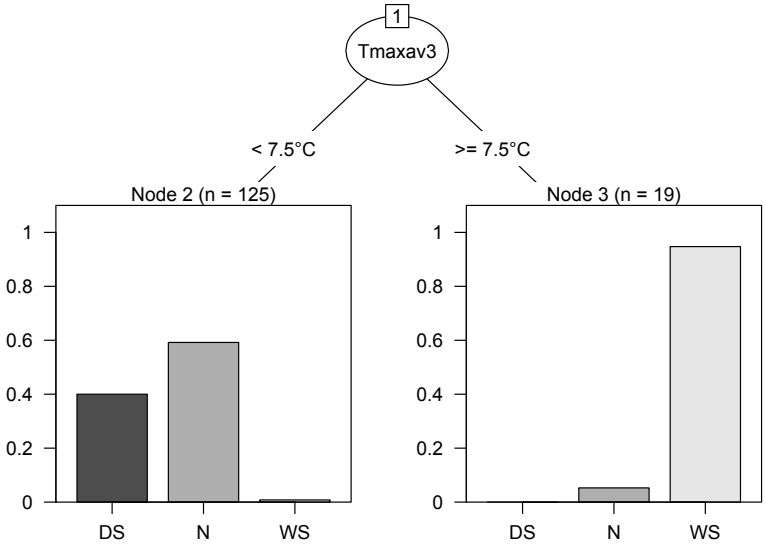



Figure 25. Cross-validated classification tree obtained when using only significantly different metrics to split days with deep wet avalanches on persistent weak layers (WS), days with deep dry slabs on persistent weak layers (DS), and days without deep slabs on persistent weak layers (N).

Table 20 Measures of model performance for cross-validated trees that split days with dry deep slabs, days with wet deep slabs, and days without deep slabs.

| Metrics used to grow tree | a | b | c | d | RPC | TSS | FAR | POD | PON | FSR |
|---|---|---|---|---|---|---|---|---|---|---|
| Cross Validated | | | | | | | | | | |
| All Daily Metrics | 59 | 24 | 17 | 52 | 0.73 | 0.46 | 0.25 | 0.68 | 0.78 | 0.32 |
| All Metrics | 59 | 24 | 17 | 52 | 0.73 | 0.46 | 0.25 | 0.68 | 0.78 | 0.32 |
| Only Significant | 74 | 51 | 1 | 18 | 0.62 | 0.25 | 0.05 | 0.26 | 0.99 | 0.74 |

4.2.4 Random Forests

The random forests show poorer model performance than the cross-validated trees for both sets of data (tables 21 & 22). The most important variable in the daily forest, based on the MDA, is minimum temperature. This is accompanied in the top ten most important variables by average maximum temperature over 7, 5, 3 and 2 days, minimum daily temperature gradient, total degrees above zero for the maximum temperature over 2, 3 and 5 days, and total degrees above zero over 5 days for the minimum temperature (Fig. 26).

Table 21 Contingency table for random forests.

| | | Observed | | | |
|---|---|---|---|---|---|
| | | Days | | Seasons | |
| Predicted | | Non Av. | Av. | Non Av. | Av. |
| | Non Av. | 49 | 24 | 13 | 7 |
| | Av. | 32 | 33 | 9 | 14 |

Table 22 Measures of model performance for random forests.

| | RPC | TSS | FAR | POD | PON | FSR |
|---|---|---|---|---|---|---|
| Days | 0.59 | 0.18 | 0.49 | 0.58 | 0.60 | 0.42 |
| Seasons | 0.63 | 0.26 | 0.39 | 0.67 | 0.59 | 0.33 |

The most important variable in the seasonal forest is total new snow in December. The next most important variables are total SWE in December, total SWE over November, December and January, and total SWE in January. These are followed by the proportion of days in November and December with conditions conducive to constructive metamorphism, and then the proportion of days in November, December and January with these conditions. Maximum snow depth in November and maximum snow depth in December are seventh and eighth in importance in the seasonal forest (Fig. 26).

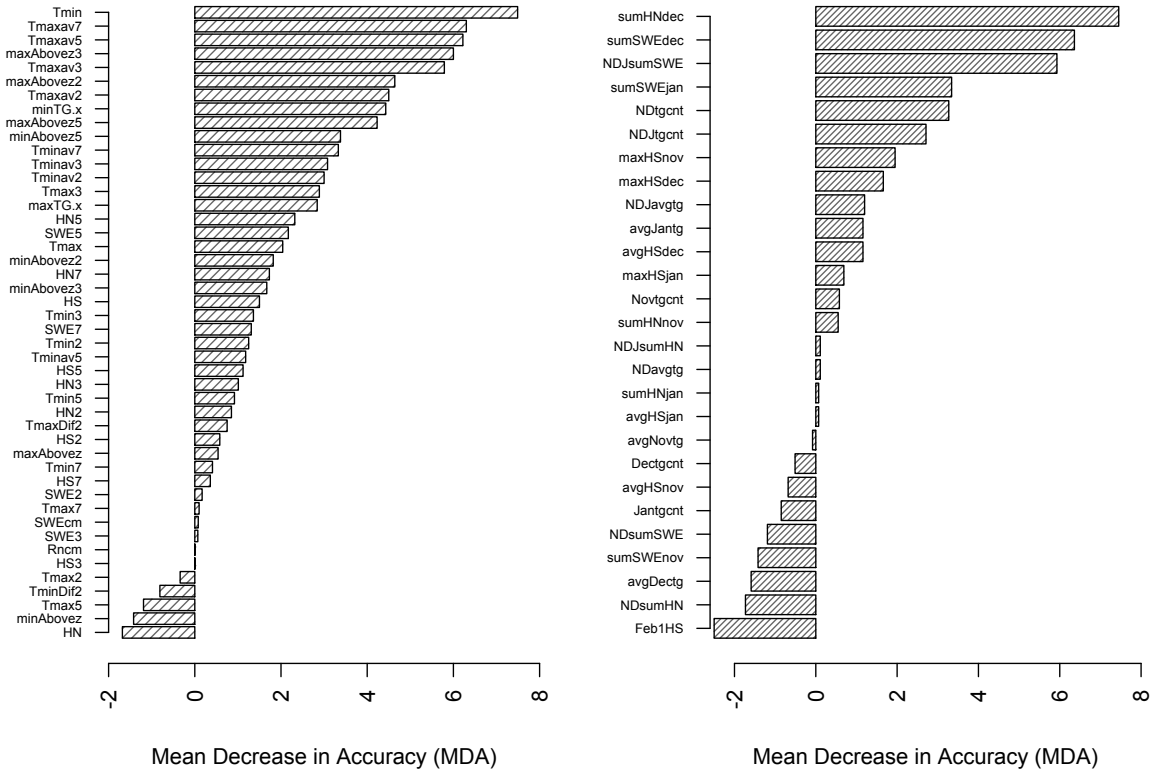



Figure 26. Variable importance shown using the MDA for each variable in the random forests for daily variables (left) and seasonal variables (right). Variables are described in tables 4,11&12.

# 5. DISCUSSION

Deep slabs on persistent weak layers late in the season were studied using records of meteorological, snowpack, and avalanche observations at Bridger Bowl. These avalanches typically have high destructive potential, and they are more difficult to predict than avalanches that release in newly fallen snow or recently buried weak layers. Forecasting these avalanches is relatively more difficult because low entropy information, such as avalanches or collapsing on deep persistent weak layers, is typically scarce or non-existent prior to these types of avalanches. Meanwhile, high entropy information that is available, such as snowpack properties and meteorological data, has increased uncertainty. If low entropy information indicating clear instability on a deep persistent weak layer does exist (i.e., observed avalanches on that weak layer), then a forecast can heavily rely on that information. However, in the absence of clear instability high entropy information must be interpreted (i.e., daily and seasonal weather summaries), and experience with local climate, local terrain, and general avalanche behavior becomes a primary factor in determining the avalanche hazard.

An avalanche forecast for any path or region begins with each winter's first snowfall (LaChapelle, 1980; McClung, 2002a). This is especially important when forecasting for deep slabs on persistent weak layers. The best way to observe a deep persistent weak layer is to watch it form, or to dig down and find it once it is buried. Ideally, both of these occur with continuous observation of the weak layer's development throughout the season. LaChapelle (1960a) mentions two general methods for forecasting dry slab avalanches. The snowpack can be analyzed, and the meteorological factors

affecting snow deposition can be analyzed. The former was regarded to be necessary when forecasting climax avalanches, while the latter was regarded as the more successful way to forecast direct-action avalanches (LaChapelle, 1960a; LaChapelle, 1965; LaChapelle, 1980).

Direct observation of the snowpack allows for slowly developing instabilities to be detected. The meteorological conditions during and after snow deposition determine the structure of the snowpack (LaChapelle, 1980), so the meteorological conditions during the time of weak layer formation may be useful information when forecasting for deeper instabilities. Interpreting weak layer properties from meteorological data has higher uncertainty than direct observation of the weak layer itself. However, direct observation is not always feasible, especially late in the season when the weak layer is deeply buried and not always spatially continuous. Therefore, this thesis used meteorological data summarized from the start of each winter, as well as meteorological data on days directly preceding deep slabs on persistent weak layers, in order to highlight meteorological metrics that could be used to aid forecasting deep slabs on persistent weak layers.

Historical records of avalanche activity and meteorological observations from Bridger Bowl ski area were used for this research (section 3.2). Avalanches that failed on deep persistent weak layers late in the season were selected through query of avalanche records (section 3.4.1), and meteorological observations were used to derive metrics that may be associated with these avalanches (section 3.4.3). Univariate analysis was performed to find meteorological metrics that were associated with days or seasons that had deep slabs on persistent weak layers, and classification trees were built to examine

meteorological metrics through multivariate analysis (section 3.4.4). In addition, meteorological and snowpack observations were reviewed in a case study of the 2011-2012 season, in which dry slab avalanches occurred on a layer of depth hoar, from December through March, and a historic wet slab avalanche cycle on the depth hoar layer culminated the season (section 4.1; and published by Marienthal et al., 2012).

Meteorological metrics that were examined in both the case study and the statistical analysis included metrics that summarized conditions over the entire season and metrics that summarized conditions up to seven days preceding deep slabs on persistent weak layers. These two sets of metrics were analyzed to answer two questions: 1) which meteorological metrics are most important to consider from the time of weak layer development (i.e., early season) when forecasting deep slabs on persistent weak layers, and 2) which meteorological metrics best distinguish between days with deep slabs on persistent weak layers and days without these avalanches? In addition to meteorological metrics, snowpack properties that may be associated with deep slabs on persistent weak layers were reviewed in the case study of the 2011-2012 season.

## 5.1 Case Study: The 2011-2012 Season

The 2011-2012 winter began with a shallow November snowpack and cold air temperatures (Fig. 13). Every day that had weather observations in November had a temperature gradient conducive to depth hoar formation (Table 15). Small, yet frequent precipitation amounts in December helped to maintain a shallow snow depth and strong temperature gradients (Fig. 13). Small depth hoar grains, resulting from these temperature gradients, were observed 33 cm deep out of a 50 cm deep snowpack, as early as

December 15th, and appeared to develop through December as larger grains were observed on December 31st and in early January (Table 14). Propagating failures in ECT tests were observed on this layer as early as December 7th, and few ECTs had propagating failures between December 8th and 24th (Fig. 14), at the same time little avalanche activity was observed on this layer (Fig. 15).

Instability on the depth hoar layer was relatively obvious in the month of December. The layer was clearly developing through December (Table 14), and stability tests and avalanches propagated after most small storms (Fig. 14 & 15). A storm with 39mm of SWE on December 30th and 31st produced widespread avalanches that failed on the depth hoar layer and stability tests had conclusive propagating failures in the depth hoar layer, through January 7th (Fig. 14 & 15a). Stability tests were in further agreement with distinct instabilities (avalanches) as tests did not show propagating failures, and avalanches did not fail on the depth hoar layer through mid-January (Fig. 14 & 15).

A continued strong temperature gradient in January appears to be one of the most significant factors that contributed to the persistence of the depth hoar layer. Characteristically cold temperatures and few small storms in January were conducive to maintaining a shallow snowpack and continued depth hoar growth (Fig. 13). Despite few detailed snowpack observations from mid-January to mid-February (Fig. 14), depth hoar grains appeared to increase in size. Grains were observed to be larger and well developed (3-5mm) in February and early March (Table 14). Snow depth was relatively far below average for November and December, and both months had a high proportion of days with temperature gradients suitable for constructive metamorphism. It may be more significant, however, that average and maximum snow depth in January, and February 1st

snow depth, were the lowest out of the 44 seasons on record (Table 15). Additionally, the proportion of days with a temperature gradient conducive to depth hoar development in November, December, and January was the highest (0.8) out of all 44 seasons. The proportion of days in January that had conditions suitable for depth hoar growth was 0.6. This was not the highest on record (0.7), but the average for this metric was 0.0 over 44 seasons (Table 15). Strong temperature gradients in a shallow snowpack appear to rarely persist through January, but when they do it may strongly suggest a red flag for late season deep slabs on persistent weak layers.

Stability test results in February and March were difficult to interpret. There were many false unstable results prior to avalanche activity in February and early March, and then potential false stable results in late March when the snowpack became too deep to affect the weak layer during stability tests. Small storms began to add up in the middle of February, increasing the overall snow depth and slab thickness (Table 14 and Fig. 13). Stability tests showed propagating failures on the depth hoar at this time (Fig. 14), yet avalanches were not triggered on this layer until late February, despite earlier additional load from storms and the application of explosives (Fig. 15). An explosive triggered avalanche (in-bounds) failed on the depth hoar layer on February 23$^{rd}$ after 10mm of SWE over 24-hrs and 84mm of SWE over seven days (Fig. 15b). A skier triggered an avalanche on the depth hoar layer out-of-bounds the next day without new snow. Two days later there was a natural avalanche near the previous one, on a deep persistent weak layer out-of-bounds, with 19mm of SWE over the preceding 48 hours and 78mm of SWE over seven days (Fig. 15b). The latter avalanche was not confirmed as sliding on depth hoar and is not marked as sliding on depth hoar on Figure 15. However, since the crown

was over 150cm deep (Fig. 15) and persistent weak layers were observed up to 100cm off the ground (Appendix B) it is significant to this study.

Precipitation continued into March with ECTs continuing to propagate failures on the depth hoar layer (Fig. 14), but avalanches were seldom observed to fail on the depth hoar layer. Although failures in ECTs became harder to initiate, they consistently produced sudden clean collapses (Appendix B). March 17th was the last day with an ECT that produced a propagating failure (Table 14). There were both natural and explosive triggered avalanches on the depth hoar beginning on March 14th after 48mm of SWE (Fig. 15c). Dry slab avalanche activity on the depth hoar layer culminated on March 20th, with two naturally triggered avalanches out-of-bounds, after 39mm of SWE over 48 hours and 75mm of SWE over the preceding seven days (Fig. 3 & 15c). Despite the difficulty of interpreting ECTs late in the season they are still useful to see the weak layer, and modified tests can reveal clean failure planes. Additionally, other evidence is available to indicate instability on deep persistent weak layers.

Cumulative snow water equivalent appears to be a good indicator of instability for deep persistent weak layers. Each dry slab avalanche cycle on the depth hoar during the 2011-2012 season occurred following more than 30mm of cumulative SWE (Fig. 15). A 48-hour storm total of 39mm of SWE was observed during the first avalanche cycle at the end of December, and could potentially be considered a minimum load required to initiate an avalanche on the depth hoar. Avalanche cycles in late February and March were preceded by over 70mm of SWE over seven days, and many had 30-50mm of SWE over the preceding 48 hours. Ultimately, using storm totals as a quantified stability test for deep persistent weak layers may be more conclusive than traditional ECT tests (e.g.,

"This layer avalanched with a load of 39mm of SWE in February" may be more valuable than "There was an ECTP11 sudden collapse on this layer"). The limitation of this would be that an avalanche must occur in order to know how much cumulative SWE preceded it.

This case study is limited by a lack of wind data. It should be noted that there is anecdotal evidence suggesting sustained high winds prior to many avalanches that fail on deep persistent weak layers (e.g., Savage, 2006; Savage and Buotte, pers. comm., 2014). High winds without snow to transport do not provide a significant load, but as more snow becomes available for transport the potential load increases. All deep slab avalanches that Savage (2006) studied had wind transport on at least one of four prior days, and 55% of deep slab avalanches occurred when four out of five prior days had wind transport (Savage, pers. comm., 2014). A combination of extreme cumulative SWE amounts and high winds greatly raises the potential for deep slabs on persistent weak layers.

Wet slabs that failed on deep persistent weak layers during the 2011-2012 season occurred when temperatures did not drop below freezing over one of the preceding two nights (Fig. 15), and the snowpack was presumably isothermal (see March 17$^{th}$ snowpit in Appendix B). After the first wet slab on depth hoar, on March 23$^{rd}$, there were wet slabs on the 24$^{th}$, 26$^{th}$, and 27$^{th}$. Only one night leading up to the 27$^{th}$ had below freezing temperatures. In addition to these above freezing temperatures, 38mm of SWE fell on the morning of the 27$^{th}$, and the historic wet slab cycle ensued. This was the only avalanche cycle on the depth hoar that had new precipitation on only one out of the preceding seven nights. However, it had a similar amount of preceding 24-48 hour SWE as previous avalanches on the depth hoar, combined with an isothermal snowpack and sustained

above freezing temperatures. Wet and dry slabs are often evaluated differently, but in this case the storm total preceding widespread wet slabs was similar to previous dry slab avalanches on this layer. The sustained above freezing temperatures was a factor that is unique to wet slabs. The weakening from sustained warmth combined with a storm similar in size to those that previously collapsed this layer created an historic wet slab cycle. The weakening from warmth may have made up for a lack in several days of cumulative load that was commonly observed before dry slabs that fail on deep persistent weak layers.

The season highlighted above presented a clear progression of meteorological conditions that often precede deep slabs on persistent weak layers. It also exemplifies the increased uncertainty of ECT stability tests on deep persistent weak layers later in the season. The persistence of a shallow snowpack and strong temperature gradients through January provided clear evidence of weak layer development and subsequently increased deep slab hazard. Snowpack observations of developing depth hoar support the validity of using summaries of meteorological metrics to imply weak layer development when direct observation of the weak layer is inconclusive or not possible. Avalanches that failed on the same depth hoar layer often occurred after similar amounts of cumulative SWE. This suggests that, while ECTs become more uncertain later in the season, daily meteorological variables (i.e., cumulative SWE) can provide evidence of quantified weak layer strength. This case study suggests that forecasting for deep slabs on persistent weak layers can be improved with the consideration of meteorological variables from the first snowfall continuously through weak layer development and conditions prior to days with potential avalanches on weak layers of concern.

## 5.2 Historical Analysis: 1968/69-2012/13

Summaries of meteorological observations over 44 seasons were statistically compared between seasons with and seasons without deep slab avalanches on persistent weak layers, and between days with and days without deep slabs on persistent weak layers. Transformation of meteorological data through summaries and statistical analysis will help reduce the relatively higher uncertainty that often accompanies the use of raw meteorological observations for an avalanche forecast (LaChapelle, 1980). Our results highlight specific meteorological patterns and trends that are common during seasons that have deep slabs on persistent weak layers, and metrics that are common on days prior to deep slabs on persistent weak layers. This will be useful by providing clear and objective distributional data to those who wish to gain more knowledge about forecasting deep slabs on persistent weak layers.

### 5.2.1 Univariate Analysis

Results of the Wilcoxon rank-sum tests show which metrics have a different distribution between seasons (or days) with and seasons (or days) without deep slabs on persistent weak layers. More than half of the seasonal metrics had significantly different central tendencies between seasons with and seasons without deep slabs on persistent weak layers (Table 16). In general, metrics with significantly different central tendencies between seasons with and seasons without deep slabs on persistent weak layers are summaries of snow depth, total precipitation, or average temperature gradient. Seasons with deep slabs on persistent weak layers generally have shallower snow depth, less total precipitation and snowfall, and stronger temperature gradients, over the entire season.

This is also generally true for these metrics summarized over the months prior to February (Table 16). Similarly, LaChapelle and Atwater (1961) mention a shallow early season snow cover in a season that had climax avalanches on depth hoar. A shallow snow depth and relatively less precipitation are conducive to a stronger temperature gradient. Including both metrics that represent relative snow depth or precipitation and calculated average temperature gradients may be somewhat redundant. However, snow depth is also important to consider in regards to determining if a weak layer, or bed surface, is higher than terrain roughness (Schweizer et al., 2009).

The only metric in November that is significantly different between seasons with and seasons without deep slabs on persistent weak layers is the proportion of days with temperature gradients conducive to constructive metamorphism. This supports the importance of early detection and observation of a persistent weak layer (LaChapelle and Atwater, 1961). This may be especially important in snow climates where depth hoar is a common weak layer as it typically forms early in the season.

In contrast, meteorological summaries over multiple months (i.e., November and December, November through January, or over the entire season) are all significantly different between seasons with and seasons without deep slabs on persistent weak layers, with the exception of temperature summaries and minimum and maximum temperature gradient over the entire season (Table 16). This supports the importance of continuous observation of conditions during weak layer development starting with the season's first snowfall (LaChapelle, 1980). In addition, the lack of a significant difference for metrics in February and March, other than average February temperature gradient, indicates that meteorological summaries prior to these months (i.e., prior to February 1$^{st}$) are sufficient

to integrate into avalanche forecasts, or statistical models, that seek to predict deep slabs on persistent weak layers at this location.

The significance of the average calculated temperature gradient in February may be representative of relatively shallower snow depths and colder temperatures that are typical during seasons that have deep slabs on persistent weak layers.

Meteorological metrics that have significantly different central tendencies between days with and days without deep slabs on persistent weak layers, and when a persistent weak layer is present and potentially active, include 24-hour rain and all summaries of temperature, except for the difference between the current and previous day's temperature (Table 17). The lack of a significant difference for loading variables (e.g., cumulative daily SWE or snowfall) highlights the difficulty in forecasting deep slab avalanches. Similar storm totals are seen prior to both days with and days without deep slabs on persistent weak layers, which leads to a high false alarm rate when relying on storm magnitude to predict deep slabs (e.g., Conlan et al., 2014).

Metrics that indicate precipitation loading prior to deep slabs may not show significant differences due to the inclusion of all seasons in the same analysis. This may be improved by using a metric that represents the amount of precipitation relative to other storms in the same season. A storm that is large enough to collapse a persistent weak layer during one season may be an insignificant addition to another season's weak layer. However, as shown in the case study (section 5.1), similar amounts of precipitation that collapse a weak layer early in the season are seen prior to avalanches on that weak layer later in the season. Additionally, the absence of wind data is limiting to this study. Any amount of precipitation that is accompanied by wind will create greater loading than just

the precipitation itself. LaChapelle and Atwater (1961) note wind combined with storms leading to avalanches on persistent weak layers, and at least one of four days prior to deep slabs studied by Savage (2006) had wind transport, and four out of five days had wind transport prior to 55% of deep slab avalanches (Savage, pers. comm., 2014).

Summaries of temperature indicate generally warmer temperatures on days with deep slabs on persistent weak layers than on days without these avalanches. This was noted in previous studies (Tracz, 2012; Conlan et al., 2014), and was explained by the fact that warmer temperatures typically accompany storms in these study areas. Warmer temperatures likely accompany deep slabs in this study due to the inclusion of wet avalanches, which are common after sustained warm temperatures (Baggi and Schweizer, 2009; Peitzsch et al., 2012).

The group of days chosen without deep slabs on persistent weak layers (section 3.4.2) is not an independent sample from the population of days that had deep slabs on persistent weak layers. The selection of a proportional number of non-event days as event days in each season suggests that there are paired groups of events and non-events that are dependent within each season. Therefore, the assumption of independent groups, which is required for the Wilcoxon rank-sum test, is violated for the univariate analysis of days with versus days without deep slabs.

It would be more appropriate to treat the groups of days with and days without deep slabs in each season as paired. However, a typical paired analysis analyzes the difference between paired observations, which are usually paired one to one. Since these groups vary in size, many of them would have to be represented by a combination of differences. This is much more involved than a typical paired analysis (e.g., paired t-test).

An alternative to using a paired analysis on these data would be to re-select a group of non-event days that were only restricted to being from seasons and months that had deep slabs, but not a proportional amount per season.

Treatment of these groups as independent when they are not provides incorrect estimates of uncertainty. This leads to inaccurate estimates of p-value, which causes some differences to be missed, or to be incorrectly supported. Therefore, our univariate analysis may incorrectly identify metrics with significantly different central tendencies between days with and days without deep slabs. Metrics with p-values closer to our cutoff of 0.05 (Table 17) will be more affected by this violation of independence.

### 5.2.2 Classification Trees for Seasons with Deep Slabs on Persistent Weak Layers

The over-fit tree that splits seasons with and seasons without deep slabs on persistent weak layers using only metrics prior to January 1st shows that seasons with less snow in December had more deep slabs on persistent weak layers (Fig. 16). Seasons with more snow in December and weaker average temperature gradients in December had more avalanches when there were a higher proportion of days in November with temperature gradients and snow depth conducive to constructive metamorphism (Fig. 16). Despite more snowfall and a relatively weak temperature gradient in December, temperature gradients in November were likely sufficient to create a weak layer that would later produce deep slabs. The seasons that had more snow in December and a stronger average temperature gradient in December did not have deep slabs on persistent weak layers. This is somewhat counterintuitive since a strong temperature gradient should imply weak layer development. However, since these were years with more

snowfall in December, the snow may have been sufficiently deep to attenuate the effect of cold temperatures deeper in the snowpack and prevent weak layer development.

The over-fit tree that uses only metrics prior to January 1st has a high POD, TSS and RPC with a low FSR. However, the cross-validation (Fig. 17) drops the RPC from 0.85 to 0.77, the TSS from 0.71 to 0.53, the POD from 0.96 to 0.58 and the FSR goes up from 0.04 to 0.42 (Table 18). Hendrikx et al. (2014) achieved a TSS of 0.47 and an RPC of 0.73 on the dataset used to fit their tree. Schweizer et al. (2009) had RPCs ranging from 0.77-0.89 and TSS ranging from 0.29-0.54 for various datasets. Despite an RPC and TSS that are comparable to previous work, the cross-validated tree shows relatively poor performance with a weak POD and FSR as these are some of the most important measures to consider for avalanche forecasting. The cross-validated trees that use the other datasets maintain comparable RPC and TSS while achieving a high POD and low FSR, at the expense of a high FAR (Table 18). These scores are more desirable and highlight the importance of evaluating meteorological conditions continuously up to the day of forecasting for deep slabs. While some metrics prior to January 1st are reasonable proxies for weak layer development, the metrics that include January observations are more robust when used to forecast deep slabs on persistent weak layers.

The high FAR for the cross-validated tree grown from datasets that use metrics through January shows that this tree is somewhat conservative, which displays the difficulty of the forecasting problem. Many seasons with low precipitation from November through January had deep slabs after February 1st, but 28% of them did not have these avalanches (Table 18). For many operations, this is likely above the acceptable level of lost operating time, or excessive closure, due to false alarms.

While the achieved model performance for the trees created to split seasons with and seasons without deep slabs on persistent weak layers is comparable to previous avalanche forecasting models, it should be noted that the sample size here is relatively low for cross-validation and growing classification trees in general (Breiman et al., 1993). Despite the small sample size, cross-validation is typically conservative in small datasets that have many predictors, which promotes the significance of the cross-validated trees' sole predictors (e.g., total SWE in November, December, and January). In addition, the maximum number of splits in the over-fit trees is limited by the stopping rules, so the smaller sample sizes are restricted to fewer splits.

The cross-validated trees that use datasets with metrics through January split seasons with and seasons without deep slabs on persistent weak layers based on the total SWE from November through January (Fig. 19). This indicates that seasons with a relatively wet early season have deep slab avalanches less frequently than those with dry early seasons. More precipitation early in the season increases the snow depth and snow density, which would decrease the overall temperature gradient and reduce weak layer development. Furthermore, abundant precipitation early in the season helps stabilize deep slab instability even if a persistent weak layer does develop. Avalanches on persistent weak layers earlier in the season may destroy persistent weak layers, and overburden pressure from load will strengthen weak layers (Brown et al., 2001).

The inclusion of January in the cumulative SWE total that splits the cross-validated tree implies the importance of continuous observation of meteorological conditions across all months. Considering the effect of November, December, and January SWE totals on late season deep slab stability has less uncertainty when all

months are combined than when they are considered individually. Stated more broadly, the effect of early season meteorological conditions on late season deep slab avalanche stability is dependent on the continuous interaction of various meteorological conditions throughout the entire early season. Observations from isolated, rather than continuous, periods of time will increase uncertainty in any avalanche forecast.

In the over-fit tree that uses the dataset with metrics prior to February 2nd, the second split divides seasons with less SWE in November, December, and January based on total December SWE (Fig. 18). Seasons with a drier December had deep slabs on persistent weak layers. Seasons with a relatively wet December, despite an overall relatively dry early season, had deep slabs when the maximum November snow depth was relatively deeper than seasons without deep slabs (Fig. 18). The inclusion of maximum November snow depth suggests that while a shallow snow depth early in the season is conducive to a weak snowpack, enough depth must be maintained in order for a weak layer, or bed surface, to form higher than the terrain roughness (e.g., Schweizer et al., 2009). Associated 95% confidence intervals for estimated snow depths (section 3.3) were within a range that would not greatly affect the threshold value for maximum November snow depth that was found to split seasons with deep slabs from seasons without deep slab avalanches.

The tree created the dataset that used metrics for the entire season has the same initial split as the previous tree (using metrics prior to February 2nd), but it replaces the second two splits with just one split based on the average temperature gradient in February (Fig. 20). This split may represent seasons that are overall cold with a shallow snowpack. This implies that seasons with generally cold temperatures and a relatively

shallow snowpack often have deep slab avalanches on persistent weak layers late in the season. A drawback to this metric is that it cannot be calculated until March 1st, which is one month after this study begins to consider deep slabs on persistent weak layers. Due to this, it is not a useful metric to consider when forecasting, but it provides some insight into the factors that are associated with deep slabs on persistent weak layers (i.e., a season with overall cold temperatures and a shallow snowpack, prior to a large storm). This is somewhat similar to findings by Fitzharris and Schaerer (1980) that indicate climax avalanches occurring in winters with sustained cold and dry conditions followed by temperatures rapidly rising above freezing with heavy precipitation.

### 5.2.3 Classification Trees for Days with Deep Slabs on Persistent Weak Layers

The over-fit trees that split days with and days without deep slabs on persistent weak layers are useful for displaying the underlying structure of the data, and show which variables are most associated with deep slabs on persistent weak layers. Meteorological conditions over days prior to deep slabs on persistent weak layers were typically warmer than days without based on the first split that uses daily minimum temperature (Fig. 21). This is consistent with previous findings of warming trends prior to deep slabs (Jamieson et al., 2001; Tracz, 2012; Conlan et al., 2014). Similar to many other studies (Jamieson et al., 2001; Schweizer et al., 2009; Tracz, 2012; Conlan et al., 2014), greater prior cumulative load is associated with deep slabs, which is shown by the splits on five-day total SWE and the change in snow depth over seven days (Fig. 21). For each of these, values indicating greater loading lead to groups with mostly deep slab days.

The final splits of the over-fit tree that uses only daily metrics utilize height of snow and a positive change in snow depth over three days. These do not provide as clear of an interpretation as the other splits, but they support the importance of considering overall snow depth and cumulative load prior to deep slabs. Days with snow depth values greater than the threshold were mostly deep slab days (Fig. 21). This is somewhat of a surprise as a shallower snowpack often indicates more instability. This split highlights the different interpretation of a snow depth observation on the day of a deep slab compared to the interpretation of a snow depth observation during persistent weak layer formation. The height of snow on the day of a deep slab avalanche has the potential to be much greater than the height of snow during weak layer formation. If a large difference in snow depth does exist between these times, then it might indicate greater instability due to a shallow snowpack (weak layer formation) followed by relatively large cumulative loading. Abrupt loading on a persistent weak layer will more likely lead to triggering of deep slabs than gradual loading (e.g., Marienthal et al., 2012), but both scenarios could have the same height of snow on a potential avalanche day. This highlights the importance of considering meteorological conditions throughout months prior to deep slab avalanche potential.

The majority of deep slab days in the previously discussed groups had dry slabs that mostly followed days with relatively warmer temperatures and larger prior cumulative loads. The days that have relatively higher prior three-day average maximum temperatures are mostly days with deep wet slabs (e.g., Fig. 24). Sustained above freezing temperatures prior to wet slabs is common as it implies a variety of conditions that are conducive to wet slab initiation (e.g., Baggi and Schweizer, 2009). While longer

periods of sustained warming may be more conducive than shorter periods of sustained warming to the initiation of deep wet slabs (e.g., more time may be required before a refreeze for water to reach a deep weak layer), other variables must also be considered to assess deep wet slab potential, such as snowpack structure and previous deep dry slab activity on persistent weak layers. Nonetheless, if the potential for deep slabs on persistent weak layers exists, then sustained above freezing temperatures is evidence to suggest increasing hazard. Tracz (2012) also noted sustained above freezing temperatures prior to deep dry slabs. Above freezing temperatures affect changing slab deformation rates in regards to dry slab avalanche initiation (Tracz, 2012), while the introduction of liquid water due to above freezing temperatures is a more important factor for wet slabs (Baggi and Schweizer, 2009; Peitzsch et al., 2012).

The over-fit tree that uses the dataset with all daily and seasonal metrics (Appendix C) has the same first four splits as the over-fit tree that uses only daily metrics (Fig. 21). However, different metrics are used to split the group of days that have a change in snow depth over seven days (HS7) greater than or equal to -10.0cm (i.e., more loading). The final splits of this tree show days with greater than 48.2cm of cumulative snowfall over seven days having more deep slabs, days with a maximum temperature greater than 3.5°C over the previous seven days and more than 0.75cm of SWE over two days having more deep slabs, and days with less than 0.75cm of SWE over the previous two days having deep slabs as well (Appendix C). This is again suggestive of large cumulative loading and warmer temperatures prior to deep slab avalanches.

The cross-validated tree that results from using each of the two previously discussed datasets uses the splits on minimum temperature, maximum average

temperature over three days, and the change in snow depth over seven days (Fig. 22). The model performance measures for the tree created from the dataset with only daily metrics differ slightly from those created from all metrics (Table 19). This is due to missing values for daily metrics on some days with deep slabs that still have a value for seasonal metrics. When only daily metrics are used these observations are eliminated from the analysis. However, when seasonal metrics are used these observations are grouped based on their values for seasonal metrics that may have formed surrogate splits (i.e., splits that were pruned from over-fit trees).

The cross-validated tree shows decreased model performance from the over-fit trees, as is expected, but maintains levels of performance that are comparable to previous work (e.g., Schweizer et al., 2009; Hendrikx et al., 2014). Unfortunately, the lack of improvement over previous work limits the sole use of this tree for forecasting purposes. This tree is useful to highlight daily metrics that are associated with deep slabs on persistent weak layers, and will aid forecasters by providing clear and concise distributional data that has been derived from decades of practitioners' observation and experience.

The over-fit tree that was created using metrics that were had significantly different central tendencies between days with and days without deep slabs on persistent weak layers (Appendix C) uses various temperature metrics as those were the only metrics with significant differences (Table 17). The groups created from this tree show both days that have deep slabs after relatively cooler temperatures and days that have deep slabs after relatively warmer temperatures. Interpretation of this tree is challenging. One pattern that becomes somewhat apparent is groups with more avalanches occurring

after splits that alternate between relatively warmer and cooler temperatures. This may suggest that more rapidly changing conditions are conducive to deep slab avalanche release (e.g., Perla, 1970; Fitzharris and Schaerer, 1980 Tracz, 2012).

The over-fit tree that uses only metrics with significantly different central tendencies has two of the same splits as the cross-validated tree created from the other datasets (Appendix C). However, the cross-validated tree created using only metrics that are significantly different has only one split on minimum 24-hour temperature (Fig. 23). This tree has lower model performance. This supports the inclusion of all metrics when growing trees as some metrics become significant after others are accounted for, which improves model performance in this case.

The same metrics that were used to split days with and days without deep slabs (Fig. 22) are used in the cross-validated tree that splits days with deep slabs by type (wet or dry). This tree is the same when using datasets other than the dataset with just metrics that were significantly different between days with and without deep slabs. The order of the metrics is different with the first split based on the average maximum temperature over three days (Fig. 24). This supports the notion that this metric is included due to its reflection of processes that lead to wet slab avalanches. Minimum temperature and snow depth change over seven days have the same thresholds values as the tree grown to split days with and days without deep slabs. This is again suggestive of warmer temperatures and greater loading prior to deep slabs.

The cross-validated tree grown from the dataset that only has metrics with significantly different central tendencies between days with and without deep slabs uses just the first split (Fig. 25). Model performance measures are much better for the tree that

uses more metrics than just those that were significantly different (Table 20). This supports the inclusion of all metrics when creating trees, as some metrics will become significant after accounting for others.

For each of the three response variables (seasons, days, and type) the cross-validated trees were the same when created from only daily metrics, or all daily and seasonal metrics. This is ideal because metrics that summarize conditions after the time that deep slab forecasting begins (e.g., February 1$^{st}$) are not useful for predictive purposes since they are not yet known.

### 5.2.4 Random Forests

The variable importance plots (Fig. 26) for the random forests show the relative effect that each variable has on the forest's predictive ability. Variables that make up splits in the cross-validated classification trees also have high importance rankings in the random forests (Fig. 26). This supports the variables' significance as predictors for deep slabs on persistent weak layers as well as interpretations of their physical relationship to deep slabs. The top variables in the seasonal random forest further suggest early winter precipitation and snow depth to be important for forecasting deep slabs on persistent weak layers later in the season. The variables representing the proportion of days with temperature gradient and snow depth conducive to weak layer development were not seen in the cross-validated trees, but their high importance in the random forest also supports the importance of early season precipitation and snow depth. These variables show the importance of the upper threshold of early season snow depth (i.e., 1m) on impeding weak layer development, while the importance of maximum November and December

snow depth may represent a minimum early season snow depth requirement, as shown in the over-fit classification tree (Fig. 18).

Minimum temperature is the most important variable in the random forest for predicting days with deep slabs on persistent weak layers (Fig. 26), and it is the primary split for the daily classification trees (Figs. 21-23). Average maximum temperature over three days is the only other variable among the ten most important in the random forest that is also seen in the classification trees. The random forest displays an overall lack of importance of loading variables for forecasting days with deep slabs, but the abundance of highly important variables that summarize maximum temperatures over multiple days strengthens the importance of sustained warming and variables that reflect rate dependent processes when forecasting deep slabs. The most important loading variable is total new snow over five days, directly followed by total SWE over five days. This supports the importance of cumulative load, yet the loading variable that is in the cross-validated classification tree (change in snow depth over seven days) has very low importance. The difficulty in forecasting deep slabs based on loading from precipitation is again apparent, likely due to high false alarm rates when trying to forecast deep slabs based a threshold amount of precipitation (e.g., Conlan et al., 2014).

The overall predictive accuracy of the random forests, based on model performance measures, is poorer than the cross-validated trees (Table 22). In addition, the random forests have a large increase in FAR and FSR, and decreased TSS. This shows the random forests' decreased ability to separate seasons (or days) with deep slabs from those without deep slabs as compared to cross-validated trees. The random forest, however, is likely a more accurate depiction of the predictive power of either dataset.

Guy and Birkeland (2013) found a wide range of success rates (62-80%) and TSS (0.17-0.38) when using random forests to classify potential avalanche trigger locations based on terrain parameters. Our model performance results are comparable despite being close to the lower end of those ranges. While the cross-validated trees provide better TSS and lower FAR and FSR, the random forests likely provide a more accurate estimate of the uncertainty in the dataset.

The variables ranked lowest in importance (Fig. 26) have negative MDAs, which means average accuracy of the random forest decreases when these variables are included. Therefore, eliminating these variables may improve the overall average predictive accuracy of the forests.

### 5.2.5 Study Limitations

The use of these models, or their important parameters, for predicting deep slab avalanches is limited by their relatively low skill, inherent uncertainty in the data, and the scope of this study. The measures of model performance are too low to rely on for an avalanche forecast, but the models and their findings are still useful when combined with experience and knowledge of other forecasting tools and avalanche behavior. Direct observation of the snowpack and avalanche occurrence, and intimate knowledge of local to regional terrain and climate, are essential to consider in addition to model predictions and meteorological parameters suggested to be of high importance.

Our results provide interpretations of observed meteorological data, which often requires some transformation or processing to reduce its uncertainty. Meteorological data comprises one of three tiers of data that are available for avalanche forecasting

(LaChapelle, 1980; McClung, 2002b). Interpretations of this data should be combined with the other two types of data: 1) snowpack structure data and 2) snow mechanical data that directly indicate instability due to an applied load (e.g., avalanche occurrence), which both have inherently lower uncertainty.

The usefulness of meteorological observations in predicting snow cover properties is also limited by spatial extent. Observations will be most representative of the snow cover at the point where they are collected, and extrapolations may be made over localized terrain while considering variation by elevation and aspect. Extrapolation of data from one point to a wider area may be more appropriate over larger areas. As the size of the forecast area decreases the problem becomes more difficult and accuracy is more important (McClung, 2002b). Using meteorological data to make an assumption about the general snowpack across a larger area may be appropriate, while a more specific forecast for a small area requires lower entropy data that is more relevant at that scale (McClung, 2002b). Furthermore, the results of this study are statistically limited to Bridger Bowl ski area. An appropriate extent for their practical application would be to similar climate types and settings as Bridger Bowl ski area that experience deep slabs on persistent weak layers.

Further uncertainty is introduced to this study due to the method of selecting days with deep slabs on persistent weak layers. This process was greatly limited by the lack of documented weak layer type in records of avalanche observations. It is possible that deep slabs not on persistent weak layers were included, and that some deep slabs on persistent weak layers were not included. Great effort was made to limit these errors by thoroughly reviewing avalanches that were selected as deep slabs on persistent weak layers. We may

have also missed events that occurred late in January due to only considering deep slabs after January 31st.

The findings of this study are further limited by the scope of the data and analysis. These data were entirely observational (i.e., no random assignment to groups) and not selected from a larger population. Therefore there is no statistical inference that can be made about the causal relationship between meteorological variables and deep slabs on persistent weak layers, and no statistical inferences can be made about these phenomena occurring outside of the study area and time period. Practical inferences can be made, however, to future situations within the study area and in areas that have similar terrain and snow climates. Therefore, our findings are practically applicable to Bridger Bowl ski area and similar regions that experience a predominantly intermountain or continental snow climate with frequent persistent weak layers near the base of the snowpack.

While classification trees and random forests can be utilized or viewed as predictive models, they are used as explanatory models in this thesis. Model performance is quantified based on predictive ability, but we can view it as a relative measure of how well we understand the phenomenon (Shmueli, 2010). Meteorological variables that we used were chosen based on known theoretical relationships, rather than including many other variables in order to improve prediction. Summaries of meteorological variables are constructs (proxies) of weak layer and snowpack characteristics that relate to avalanche formation and release. Therefore, our models are explanatory due their nature of testing causal hypotheses (i.e., relationships between proxies of snowpack properties and avalanche occurrence) (Shmueli, 2010). Meteorological summaries represent snowpack processes (e.g., weak layer and slab formation, applied loads, weakening) and

characteristics (e.g., snowpack structure, isothermality) that are usually not documented on a consistent basis, while meteorological observations are documented regularly at most operations.

# 6. CONCLUSION

Deep slab avalanches on persistent weak layers are difficult to forecast largely due to a common lack of (low entropy) data that indicates obvious instability directly preceding their release. Other data is available to aid forecasting deep slabs, but it often requires some processing and experience in order to maximize its usefulness. Meteorological data can be used to assess deep slab instability, despite inherent higher entropy. Summaries of meteorological data that are presented in this thesis may be useful to minimize uncertainty when forecasting deep slabs on persistent weak layers. These summaries (metrics) highlight seasonal and daily meteorological trends that are associated with deep slabs, which are typically less recognizable given minimal (or zero) experience.

Records of avalanche, meteorological, and snowpack observations from Bridger Bowl ski area in SW Montana were used to examine meteorological metrics associated with deep slabs on persistent weak layers. This was achieved by conducting a case study of one season that had historic deep slabs on persistent weak layers, and by performing a statistical analysis of historical data from 44 seasons. Univariate analysis, classification trees, and random forests were used to statistically examine the historical data to find meteorological metrics that are common during seasons with deep slabs on persistent weak layers and metrics that are common prior to days with deep slabs on persistent weak layers.

The 2011-2012 season demonstrated the increasing difficulty of forecasting avalanches on persistent weak layers as the season progresses. Obvious clues of

instability on the depth hoar became more scarce as there were less avalanches on the depth hoar and ECT test results became harder to interpret and less agreeable with avalanche activity later in the season. The univariate analysis highlights further difficulty in forecasting deep slabs by showing no significant difference in metrics that represent loading between days with and days without deep slabs on persistent weak layers. While common data (e.g., stability tests and avalanche occurrence) often becomes less obvious regarding deep slab instability, other available data may highlight patterns that indicate instability.

The univariate analysis, classification trees, and random forests all suggest that seasons with deep slabs on persistent weak layers had overall less precipitation and a shallow snowpack in November, December, and January. While a drier early season is common, a November snow depth that is deep enough to create a sufficient weak layer is also common to seasons with deep slabs on persistent weak layers. Our results also support the importance of considering continuous meteorological metrics through Janauary (e.g., cumulative SWE), rather than metrics from isolated months or days (e.g., February 1$^{st}$ snow depth), to forecast deep slabs on persistent weak layers.

A continuous shallow snowpack and cold air temperatures from November 2011 through January 2012, which coincided with observed depth hoar growth, supports the continuous observation of meteorological metrics as a proxy for weak layer development. The use of meteorological metrics that summarize conditions during weak layer development is further supported by the results from the univariate analysis that indicate metrics after February 1$^{st}$, and metrics summarized in individual months, to be generally insignificant. The results from classification trees that use metrics prior to February 2$^{nd}$ to

adequately predict (relative to this and previous studies) seasons with deep slabs on persistent weak layers also support the strength of continuous observation.

Days with deep slabs on persistent weak layers were associated with warmer temperatures and greater prior cumulative precipitation. In addition to general cumulative precipitation, relative SWE amounts prior to previous avalanches on a persistent weak layer may be useful to indicate the strength of that layer and the possible load required for it to collapse. Similar precipitation totals were observed prior to all storms that triggered deep dry slabs on persistent weak layers in the 2011-2012 season. The historic deep wet slabs in the 2011-2012 season had similar 24-48 hour SWE totals as deep dry slabs during that season, but there was no precipitation between two and seven days before the historic wet slab cycle. Instead there was sustained above freezing temperatures to weaken the snowpack prior to stress from the short, yet significant, storm. These wet slabs were also preceded by natural wet slabs on the persistent weak layer, which failed after sustained above freezing temperatures in a snowpack that was unaltered by ski area activities.

Both dry and wet deep slabs had prior warming, but the mechanisms that affect instability are different for each. Wet slabs are common after sustained above freezing temperatures that weaken the snowpack. Dry deep slabs are common after a warming trend, which may alter snowpack properties, and sometimes accompanies precipitation. Although deep slabs are shown to be more often associated with warmer temperatures, they were also seen after extreme drops in temperature during the 2011-2012 season, and further statistical evidence suggests the overall importance of rate dependent variables (e.g., Perla, 1970). The association of changes in air temperature (up or down), rather

than just warmer temperatures, with deep slabs is supported by the significant difference of most temperature metrics between days with and days without deep slabs on persistent weak layers, and the apparent switch between warmer and colder temperatures that lead to days with deep slabs that is shown in the over-fit classification tree grown from the metrics with significantly different central tendencies.

Seasonal variables associated with deep slabs are best viewed as rate dependent as well. Evidence from classification trees and the case study suggest that seasonal snow depth during seasons with deep slabs has a slow rate of accumulation followed by a rapid increase in the rate of accumulation. This further suggests that a season with a consistent rate of snow accumulation would less likely have deep slabs as the snowpack gradually adjusts to loading. These seasons would have either very few avalanches overall, or frequent smaller avalanches. Varying rates of snow accumulation throughout the season may indicate greater deep slab instability.

Cross-validated classification trees showed an overall improvement in model performance when more metrics were included as explanatory variables. Using only metrics that are significantly different in a univariate analysis limits the trees' predictive ability. Metrics that are not significant on their own can become significant after accounting for other metrics in a multivariate analysis. All cross-validated trees that were grown with only significant metrics were pruned to one split and had lower model performance than other trees. Overall model performance of cross-validated trees is comparable to previous work, which supports the use of these results as supplemental information to forecast deep slabs on persistent weak layers. The cross-validated trees have a high FAR and many have a high FSR making them unsuitable solely for

predicting deep slabs. However, the metrics used in these trees are useful to consider when forecasting deep slabs on persistent weak layers.

This study is limited by spatial and temporal scale, inherent uncertainty in the data, and the method of deep slab avalanche selection and definition. While the results of this study highlight meteorological patterns and trends to look for when forecasting deep slabs, the findings must be considered in combination with a general understanding of avalanche dynamics, knowledge of local terrain and climate, and the seasonal snowpack and avalanche history.

# REFERENCES CITED

Akitaya, E., 1974. Studies on Depth Hoar. Contributions from the Institute of Low Temperature Science, v. 26, pp. 1-67.

Armstrong, R.L., 1985. Metamorphism in a subfreezing, seasonal snow cover: The role of thermal and water vapor pressure conditions. PhD dissertation. Department of Geography, University of Colorado, Boulder, CO, USA, p. 175.

Armstrong, R.L., and Brun, E., 2008. Snow and Climate: Physical Processes, Surface Energy Exchange and Modeling. Cambridge University Press. p. 256.

Atwater, M.M., 1954. Snow avalanches. Scientific American, v. 190, pp. 26-31.

Arons, E.M., Colbeck, S.C., and Gray, J.M.N.T., 1998. Depth hoar growth rates near a rocky outcrop. Journal of Glaciology, v. 44(148), pp. 477-484.

Baggi, S. and Schweizer, J., 2009. Characteristics of wet-snow avalanche activity: 20 years of observations from a high alpine valley (Dischma, Switzerland). Natural Hazards, v. 50(1), pp. 97-108.

Birkeland, K., 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, v. 30, pp.193-199.

Birkeland, K. W., Johnson, R.F., and Schmidt, D.S., 1998. Near-surface faceted crystals formed by diurnal recrystallization: a study of weak layer formation in the mountain snowpack and its contribution to snow avalanches. Arctic and Alpine Research, v. 30(2), pp.200-204.

Bois, P., Obled, C., and Good, W., 1975. Multivariate data analysis as a tool for day-by-day avalanche forecast. Proceedings of Symposium on Snow Mechanics, Grindlewald, Switzerland, IAHS Publ. No. 114, pp. 391-403.

Bovis, M.J., 1977. Statistical forecasting of snow avalanches, San Juan Mountains, southern Colorado, USA. Journal of Glaciology, v. 18(78), pp. 87-99.

Bradley, C.C., 1970. The location and timing of deep slab avalanches. Journal of Glaciology, v. 9(56), pp. 253-261.

Bradley, C.C., and Bowles, D., 1967. Strength–load ratio, an index of deep slab avalanche conditions. Physics of Snow and Ice, Institute of Low Temperature Science, Hokkaido University, Japan, 1(2), pp. 1243–1253.

Bradley, C.C., Brown, R.L., and Williams, T.T., 1977. Gradient metamorphism, zonal weakening of the snow-pack and avalanche initiation. Journal of Glaciology, v. 19(81), pp. 335-342.

Breiman, L., 2001. Random forests. Machine Learning, 45 (1), pp. 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J., 1993. Classification and Regression Trees. Chapman and Hall, New York, USA, p. 358.

Brown, R., Satyawali, K., Lehning, M., and Bartelt, P., 2001. Modeling the changes in microstructure of snow during metamorphism. Cold Regions Science and Technology, v. 33(2–3), pp. 91–101.

Buser, O., 1983. Avalanche forecast with the method of nearest neighbours: an interactive approach. Cold Regions Science and Technology, v. 8(2), pp. 155–163.

Buser, O., 1989. Two years experience of operational avalanche forecasting using nearest neighbours method. Annals of Glaciology, v. 13, pp. 31–34.

CAIC, 2014. Colorado Avalanche Information Center: Statistics and Reporting. Retrieved September, 2014. URL: http://avalanche.state.co.us/accidents/statistics-and-reporting/

Carvelli, P., 2008, Boot packing and "systematic application of explosives": Shear plane disruption technique in the continental climate. Proceedings of the 2008 International Snow Science Workshop, Whistler, BC, pp. 337-344.

Chabot, D., 2002. Avalanche education for snowmobiles: efforts of The Gallatin National Forest Avalanche Center. Proceedings of the 2002 International Snow Science Workshop, Penticton, B.C., pp. 237-241.

Chabot, D., Staples, M., and Knoff, E., 2012. 2011-2012 Annual Report, Gallatin National Forest Avalanche Center, United States Department of Agriculture, p. 28.

Clark, M., 2013. General additive models: getting started with additive models in R. Center for Social Research, University of Notre Dame, p. 30.

Colbeck, S.C., Akitaya, E., Armstrong, R., Gubler, H., Lafeuille, J., Lied, K., McClung, D., and Morris, E., 1990. The international classification of seasonal snow on the ground. International Commission on Snow and Ice (ICSI). International Association of Scientific Hydrology, Wallingford, Oxon, U.K. p. 23.

Comey, B., and McCollister, C., 2008. Deep slab instability characterizing the phenomena - part 1. Proceedings of the 2008 International Snow Science Workshop, Whistler, BC, pp. 315–321.

Conlan, M.J.W., Tracz, D.R., and Jamieson, B., 2014. Measurements and weather observations at persistent deep slab avalanches. Cold Regions Science and Technology, v. 97, pp. 104-112. doi:10.1016/j.coldregions.2013.06.011

Conway, H., Carran, W., and Carran, A., 2009. Rain-on-snow avalanches: forecasting return to stability. International Snow Science Workshop ISSW, Davos, Switzerland, 27 September - 2 October 2009. Swiss Federal Institute for Forest, Snow and Landscape Research WSL, pp. 267-270.

Conway, H., and C.F. Raymond, 1993. Snow stability during rain. Journal of Glaciology, v. 39(133), pp. 635-642.

Conway, H., and Wilbour, C., 1999. Evolution of snow slope stability during storms. Cold Regions Science and Technology, v. 30, pp. 67-77.

Davis, R.E., Elder, K., Howlett, D., and Bouzaglou, E., 1999. Relating storm and weather factors to dry slab avalanche activity at Alta, Utah, and Mammoth Mountain, California, using classification and regression trees. Cold Regions Science and Technology, v. 30(1–3), pp. 79–89.

Delparte, D., Jamieson, B., and Waters, N., 2008. Statistical runout modeling of snow avalanches using GIS in Glacier National Park, Canada. Cold Regions Science and Technology, v. 54, pp. 183-192.

Doswell, C.A., Davies-Jones, R., Keller, D.L., 1990. On summary measures of skill in rare event forecasting based on contingency tables. Weather and Forecasting, v. 5 pp. 576–585.

Fitzharris, B., and Schaerer, P., 1980. Frequency of major avalanche winters. Journal of Glaciology, v. 26(94), pp. 43-52.

Floyer, J.A., and McClung, D.M., 2003. Numerical avalanche prediction: Bear Pass, British Columbia, Canada. Cold Regions Science and Technology, v. 37, pp. 333-342.

Föhn, P.M.B., 1993. Characteristics of weak snow layers or interfaces. Proceedings of the 1992 International Snow Science Workshop, Breckenridge, Colorado, USA, pp. 160–170.

Föhn, P., Good, W., Bois, P., and Obled, C., 1977. Evaluation and comparison of statistical and conventional methods of forecasting avalanche hazard. Journal of Glaciology, v. 19(81), pp. 375-387.

Föhn, P., Stoffel, M., and Bartelt, P., 2002. Formation and forecasting of large (catastrophic) new snow avalanches. Proceedings of the 2002 International Snow Science Workshop, Penticton, B.C., pp. 141-148.

Gallagher, D., 1967. The snowy torrents: avalanche accidents in the United States 1910-1966. Alta Avalanche Study Center, U.S. Department of Agriculture, p. 144.

Giddings, J.C., and LaChapelle, E., 1962. The formation rate of depth hoar. Journal of Geophysical Research, v. 67(6), pp. 2377-2383.

Gleason, J.A., 1995. Terrain parameters in avalanche starting zones and their effect on avalanche frequency. Proceedings of the 1994 International Snow Science Workshop, Snowbird, Utah, U.S.A., pp. 393-404.

Gray, D.M., and Male, D.H., 1981. Handbook of Snow: Principles, Processes, Management and Use. Blackburn Press, p. 776.

Greene, E., Atkins, D., Birkeland, K., Elder, K., Landry, C., Lazar, B., McCammon, I., Moore, M., Sharaf, D., Sterbenz, C., Tremper, B., Williams, K., 2010. Snow, Weather, and Avalanches: Observational Guidelines for Avalanche Programs in the United States. American Avalanche Association, Pagosa Springs CO, p. 152.

Guy, Z.M., and Birkeland, K.W., 2013. Relating complex terrain to potential avalanche trigger locations. Cold Regions Science and Technology, v. 86, pp. 1-13.

Harvey, S. and Heierli, J., 2009. Interpretation of prevalent avalanche scenarios on the basis of the anticrack model. International Snow Science Workshop ISSW, Davos, Switzerland, 27 September - 2 October 2009. Swiss Federal Institute for Forest, Snow and Landscape Research WSL, pp. 276-281.

Heierli, J., and Zaiser, M., 2008. Failure initiation in snow stratifications containing weak layers: nucleation of whumpfs and slab avalanches. Cold Regions Science and Technology, v. 52 (3), pp. 385–400.

Heierli, J., Gumbsch, P., and Zaiser, M., 2008. Anticrack nucleation as triggering mechanism for slab avalanches. Science, v. 321 (5886), pp. 240–243.

Heierli, J., Birkelnad, K.W., Simenhois, R., and Gumbsch, P., 2011. Anticrack model for skier triggering of slab avalanches. Cold Regions Science and Technology, v. 65, pp. 372-381. doi:10.1016/j.coldregions.2010.10.008

Hendrikx, J., Owens, I., Carran, W., and Carran, A., 2005. Avalanche activity in an extreme maritime climate: the application of classification trees for forecasting. Cold Regions Science and Technology, v. 43, pp. 104–116. doi: http://dx.doi.org/10.1016/j.coldregions.2005.05.006

Hendrikx, J., Murphy, M., and Onslow, T., 2014. Classification trees as a tool for operational avalanche forecasting on the Seward Highway, Alaska. Cold Regions Science and Technology, v. 97, pp. 113-120. doi: http://dx.doi.org/10.1016/j.coldregions.2013.08.009

Hothorn, T., and Zeileis, A., 2014. partykit: A Modular Toolkit for Recursive Partytioning in R. *Working Papers in Economics and Statistics, Research Platform Empirical and Experimental Economics*, University of Innsbruck, 14 pp. URL: http://EconPapers.RePEc.org/RePEc:inn:wpaper:2014-10

James, G., Witten, D., Hastie, T., and Tibshirani, T., 2013. An Introduction to Statistical Learning. Springer, New York, p. 426.

Jamieson, J., and Johnston, C., 1992. Snowpack characteristics associated with avalanche accidents. Canadian Geotechnical Journal, v. 29(5), pp. 862-866.

Jamieson, B., Geldsetzer, T. and Stethem, C., 2001. Forecasting for deep slab avalanches. Cold Regions Science and Technology, 33(2-3), pp. 275-290.

Jones, A.S.T., and Jamieson, B., 2001. Meteorological forecasting variables associated with skier-triggered dry slab avalanches. Cold Regions Science and Technology, v. 33, pp. 223-236.

Judson, A., and Erickson, B.J., 1973. Predicting avalanche intensity from weather data: a statistical analysis. U.S. Department of Agriculture, Forest Service Research Paper, Fort Collins, CO., p. 12.

Judson, A., and Doesken, N., 1999. Density of freshly fallen snow in the central Rocky Mountains. Climatology Report No. 99-2, U.S. Department of Interior, Denver, CO., p. 22.

Kattelmann, R., 1984. Wet slab instability. Proceedings of the International Snow Science Workshop, Aspen, CO, U.S.A., pp. 102–108.

LaChapelle, E., 1962. The density distribution of new snow. Alta Avalanche Study Center, p. 6.

LaChapelle, E., 1960a. Snow avalanches: their characteristics, forecasting and control. U.S. Dept. of Agriculture, Forest Service, Wasatch National Forest, p. 8.

LaChapelle, E., 1960b. Critique in heat and vapor transfer in snow. Alta Avalanche Study Center, p. 12.

LaChapelle, E., 1965. Avalanche forecasting – a modern synthesis. U.S. Dept. of Agriculture, Forest Service, Wasatch National Forest, p. 9.
LaChapelle, E. R.,1980. The fundamental processes in conventional avalanche forecasting. Journal of Glaciology, v. 26(94), pp. 75–84.

LaChapelle, E. R., 1985. The ABC of Avalanche Safety, 2nd ed. The Mountaineers Books, Seattle, WA, USA, p. 112.

LaChapelle, E., and Atwater, M.M., 1961. The climax avalanche. U.S. Department of Agriculture, Forest Service, p. 34.

Liaw, A., and Wiener, M., 2002. Classification and regression by randomForest. R News. 2(3), pp. 18-22.

Logan, S., Birkeland, K., Kronholm, K., and Hansen, K., 2007. Temporal changes in the slope-scale spatial variability of the shear strength of buried surface hoar layers. Cold Regions Science and Technology, v. 47, pp. 148-158.

Marbouty, D., 1980. An experimental study of temperature-gradient metamorphism. Journal of Glaciology, v. 26(94), pp. 303-312.

Marienthal, A., Hendrikx, J., Chabot, D., Maleski, P., and Birkeland, K., 2012. Depth hoar, avalanches, and wet slabs: a case study of the historic March, 2012 wet slab avalanche cycle at Bridger Bowl, Montana. Proceedings of the 2012 International Snow Science Workshop, Anchorage, AK. pp. 62-68.

McClung, D.M., 1979. Shear fracture precipitated by strain softening as a mechanism of dry slab avalanche release. Journal of Geophysical Research, v. 84, pp. 3519–3526.

McClung, D.M., 1981. Fracture mechanical models of dry slab avalanche release. Journal of Geophysical Research, v. 86, pp. 10783-10790.

McClung, D.M., 1996. Effects of temperature on fracture in dry slab avalanche release, Journal of Geophysical Research, v. 101, pp. 21907-21920.

McClung, D. M., 2002a. The elements of applied avalanche forecasting Part I: The human issues. Natural Hazards, v. 26, pp. 111–129.

McClung, D. M., 2002b. The elements of applied avalanche forecasting Part II: Physical issues and the rules of applied avalanche forecasting. Natural Hazards, v. 26, pp. 131–146.

McClung, D.M., 2005. Approximate estimates of fracture speeds for dry slab avalanches. Geophysical Research Letters, v. 32, L08406. doi:10.1029/ 2005GL022391.

McClung, D.M., and Tweedy, J., 1994. Numerical avalanche prediction: Kootenay Pass, British Columbia, Canada. Journal of Glaciology, v. 40(135), pp. 350-358.

McClung, D.M., and Schweizer, J., 1996. Effect of snow temperature on skier triggering of dry snow slab avalanches. Proceedings of the 1996 International Snow Science Workshop, Banff, Alberta, Canada, pp. 113-117.

McClung, D., and Schaerer, P., 2006. The Avalanche Handbook, third ed. The Mountaineers Books, Seattle, WA, USA, p. 342.

McCollister, C., Birkeland, K., Hansen, K., Aspinall, R., and Comey, R., 2003. Exploring multi-scale spatial patterns in historical avalanche data, Jackson Hole Mountain Resort, Wyoming. Cold Regions Science and Technology, v. 37, pp. 299-313.

McMannis, W.J., 1955. Geology of the Bridger Range, Montana. Bulletin of the Geological Society of America, v. 66, pp.1385–1430.

Mears, A.I., 1989. Regional comparisons of avalanche-profile and runout distance. Arctic and Alpine Research, v. 21(3), pp. 283-287.

Meister, R., 1991. Grundlagen der schweizerischen Lawinenwarnung. Vermessung, Photogrammetrie, Kulturtechnik, v. 3(91), pp. 8895.

Mock, C.J. and Birkeland, K.W., 2000. Snow avalanche climatology of the western United States mountain ranges. Bulletin of the American Meteorological Society, v. 81(10), pp. 2367-2392.

Obled, C., and Good, W., 1980. Recent developments of avalanche forecasting by discriminant analysis techniques: a methodological review and some applications to the Parsenn area (Davos, Switzerland). Journal of Glaciology, v. 25(92), pp. 315–346.

Perla, R., 1970. On contributory factors in avalanche hazard evaluation. Canadian Geotechnical Journal, v. 7(4), pp. 414-419.

Perla, R., 1977. Slab avalanche measurements. Canadian Geotechnical Journal, v. 14, pp. 206-213.

Peitzsch, E.H., Hendrikx, J., Fagre, D.B., and Reardon, B., 2012. Examining spring wet slab and glide avalanche occurrence along the Going-to-the-Sun Road corridor, Glacier National Park, Montana, USA. Cold Regions Science and Technology, v. 78, pp. 73-81.

Pfeffer, W.T., and Mrugala, R., 2002. Temperature gradient and initial snow density as controlling factors in the formation and structure of hard depth hoar. Journal of Glaciology, v. 48(163), pp. 485-494.

Ramsey, F.L., and Schafer, D.W., 2002. The Statistical Sleuth: A Course in Methods of Data Analysis, second ed. Duxbery, Pacific Grove, CA, USA, p. 742.

R Core Team, 2013. R. a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: http://www.R-project.org/.

Reardon, B., and Lundy, C., 2004. Forecasting for natural avalanches during spring opening of the Going-to-the-Sun Road, Glacier National Park, USA. Proceedings of the international Snow Science Workshop, Jackson Hole, WY, pp. 2367–2392.

Romig, J., 2004. March Wet Avalanche Prediction at Bridger Bowl ski area, Montana. M.S. Thesis. Department of Earth Sciences, Montana State University, Bozeman, MT, USA, p. 297.

Sahn, K.R., 2010. Avalanche risk reduction in the continental climate: how to implement an effective boot packing program. Proceedings of the 2010 International Snow Science Workshop, Squaw Valley, CA, pp. 296-301.

Savage, S., 2006. Deep slab avalanche hazard forecasting and mitigation: the south face at Big Sky ski area. Proceedings of the 2006 International Snow Science Workshop, Telluride, CO, pp. 483–490.

Schweizer, J., and Föhn, P.M., 1996. Avalanche forecasting: an expert system approach. Journal of Glaciololgy, v. 42 (141), pp. 318–332.

Schweizer, J., and Jamieson, J.B., 2001. Snow cover properties for skier triggering of avalanches. Cold Regions Science and Technology, v. 33(2-3), pp. 207-221.

Schweizer, J., and Jamieson, J.B., 2003. Snowpack properties for snow profile analysis. Cold Regions Science and Technology, v. 37, pp. 233-241.

Schweizer, J., and Jamieson, J.B., 2010. On surface warming and snow instability. Proceedings of the 2010 International Snow Science Workshop, Squaw Valley, CA, pp. 619-622.

Schweizer, J., Jamieson, J.B., and Schneebeli, M., 2003. Snow avalanche formation. Reviews of Geophysics, v. 41(4), pp. 1016-1041.

Schweizer, J., and Lütschg, M., 2001. Characteristics of human-triggered avalanches. Cold Regions Science and Technology, v. 33(2-3), pp. 147-162.

Schweizer, J., Mitterer, C., and Stoffel, L., 2009. On forecasting large and infrequent snow avalanches. Cold Regions Science and Technology, v. 59, pp. 234-241.

Shmueli, G., 2010. To explain or to predict. Statistical Science, v. 25, pp. 289-310.

Skipp, B., Lageson, D.R., and McMannis, W.J., 1999. Geologic map of the Sedan quadrangle, Gallatin and Park Counties, Montana. U.S. Geological Survey Geologic Investigations Series I-2634, 1 plate, scale 1:48,000. URL: http://pubs.usgs.gov/imap/i-2634/index.html, accessed June 11, 2014.

Sturm, M., Holmgren, J., König, M., and Morris, K., 1997. The thermal conductivity of seasonal snow. Journal of Glaciology, v. 43 (143), pp. 26-41.

Tracz, D., 2012. Deep Snow Slab Avalanches. M.S. Thesis. Department of Civil Engineering, University of Calgary, Alberta p. 214.

Tremper, B., 2008. Staying Alive in Avalanche Terrain, second ed. The Mountaineers Books, Seattle, WA, USA, p. 318.

Therneau, T., Atkinson, B., and Ripley, B., 2013. rpart: Recursive Partitioning and Regression Trees. *R package version 4.1-5*. URL: http://CRAN.R-project.org/package=rpart

van Herwijnen, A., and Heierli, J., 2009. Measurement of crack-face friction in collapsed weak snow layers. Geophysical Research Letters, v. 36, L23502.

Voight, B., Armstrong, B.R., Armstrong R.L., Bowles D., Brown R.L., Ferguson S.A., Fredston, J.A., Kiusalaas, McFarlane, R.C., and Penniman, R., 1990. Snow avalanche hazards and mitigation in the United States. National Academy Press, Washington, DC, p. 84.

Wallace, J.M., and Hobbs, P.V., 2006. Atmospheric Science, second ed. Academic Press, Burilington, MA, p. 483.

Wieland, M., Hendrikx, J., and Birkeland, K., 2012. The effectiveness of boot packing for snowpack stabilization. Proceedings of the 2012 International Snow Science Workshop, Anchorage, AK, pp. 993-997.

Williams, K., 1994. The U.S. Forest Service – Westwide Avalanche Network. Proceedings of the 1994 International Snow Science Workshop, Snowbird, UT, pp. 644. URL: http://arc.lib.montana.edu/snow-science/item/1385

Wilks, D., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press, p. 467.

Wilson, A., Schweizer, J., Johnston, C.D., and Jamieson, J.B., 1999. Effects of surface warming of a dry snowpack, Cold Regions Science and Technology, v. 30, pp. 59-65.

Zuur A.F., Ieno, E.N., Walker, N.J., Saveliev A.A., and Smith, G.M., 2006. Mixed Effects Models and Extensions in Ecology with R. Springer, p. 574.

# APPENDICES

# APPENDIX A

## MODEL RESIDUALS FOR IMPUTATIONS

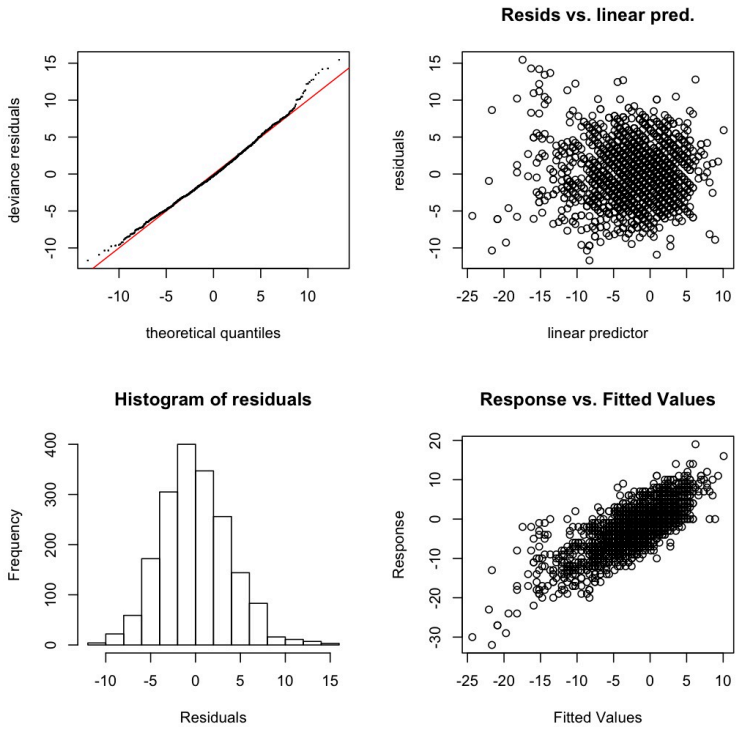



Figure A 1. Residual plots for a linear model of maximum temperature.

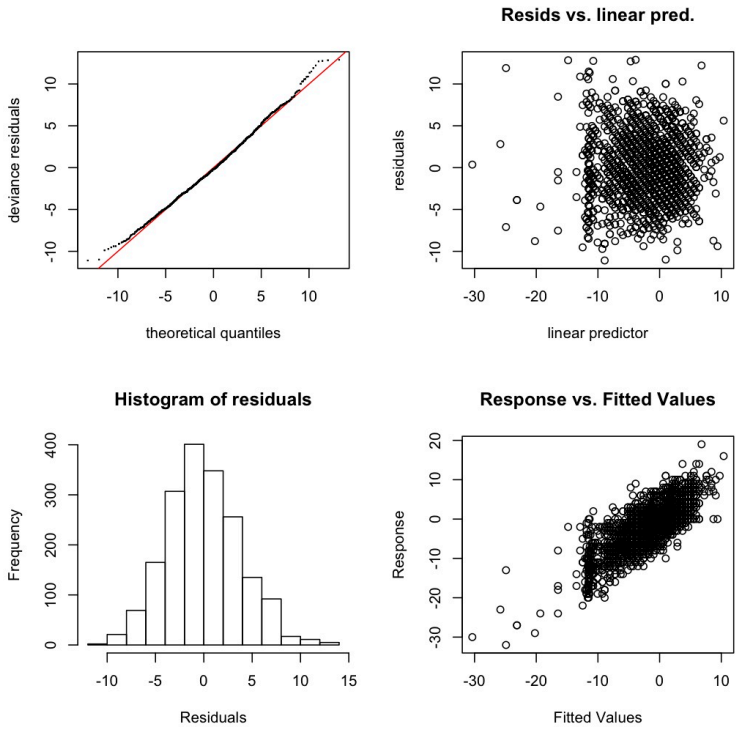



Figure A 2. Residual plots for a model with a smoothing function for maximum temperature.

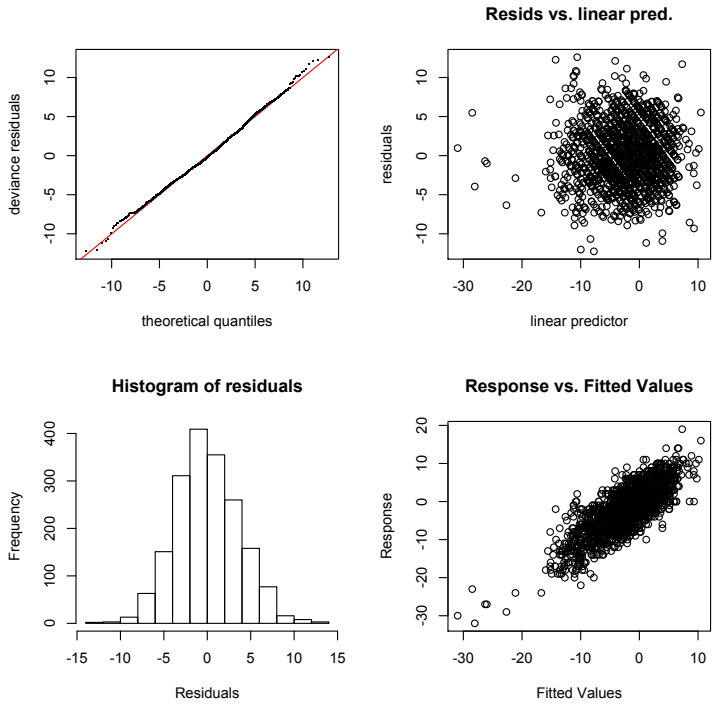



Figure 22.5. Residuals from final model for maximum temperature.

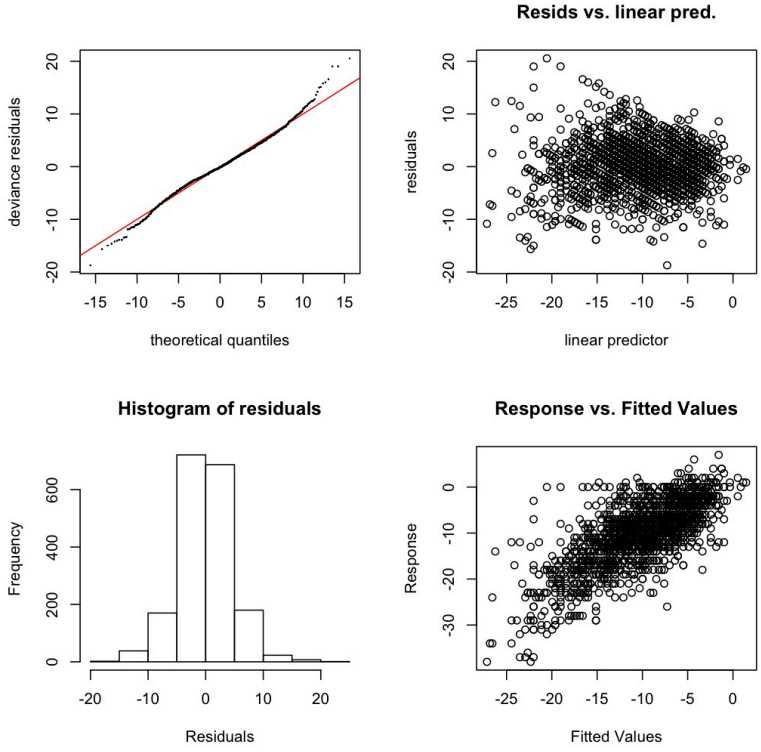



Figure A 3. Residual plots for a linear model of minimum temperature.

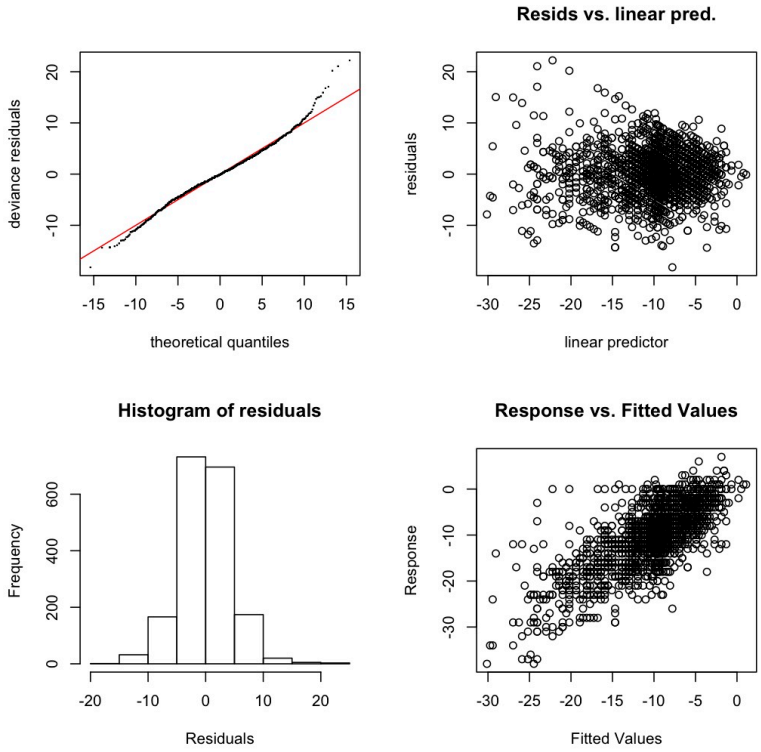



Figure A 4. Residual plots for a model with a smoothing function for minimum.

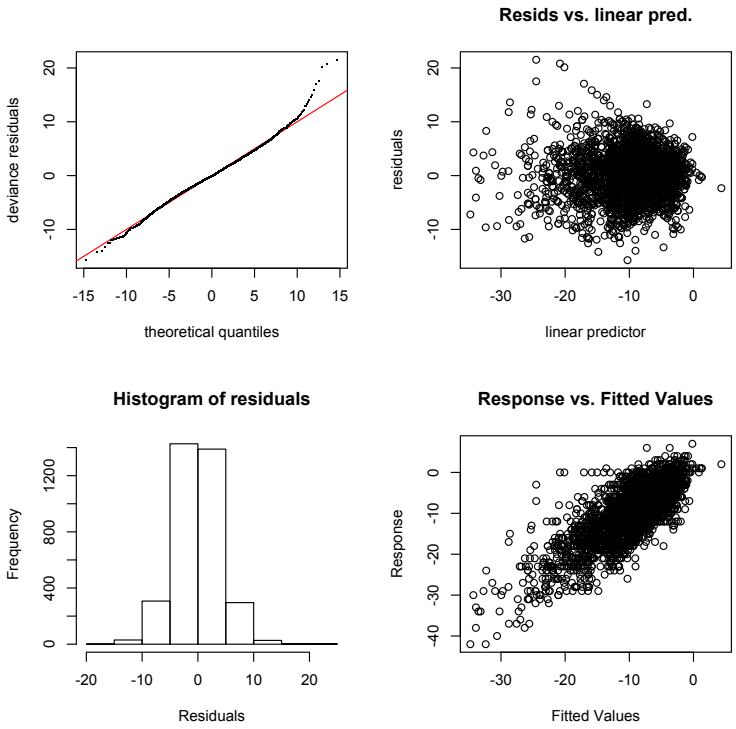



Figure A 5. Residuals from final model for minimum temperature.

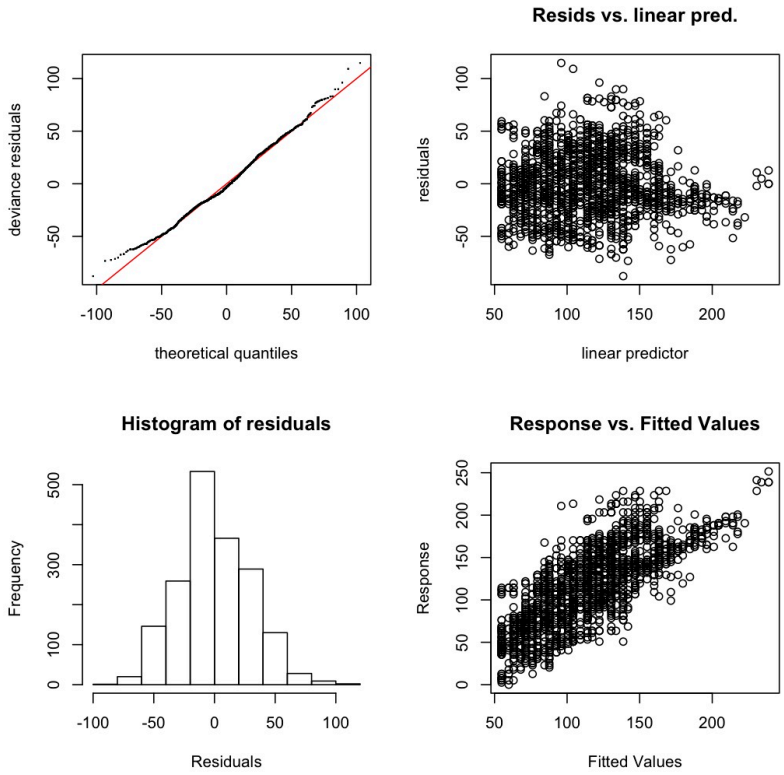



Figure A 6. Residual plots for a linear model for snow depth.

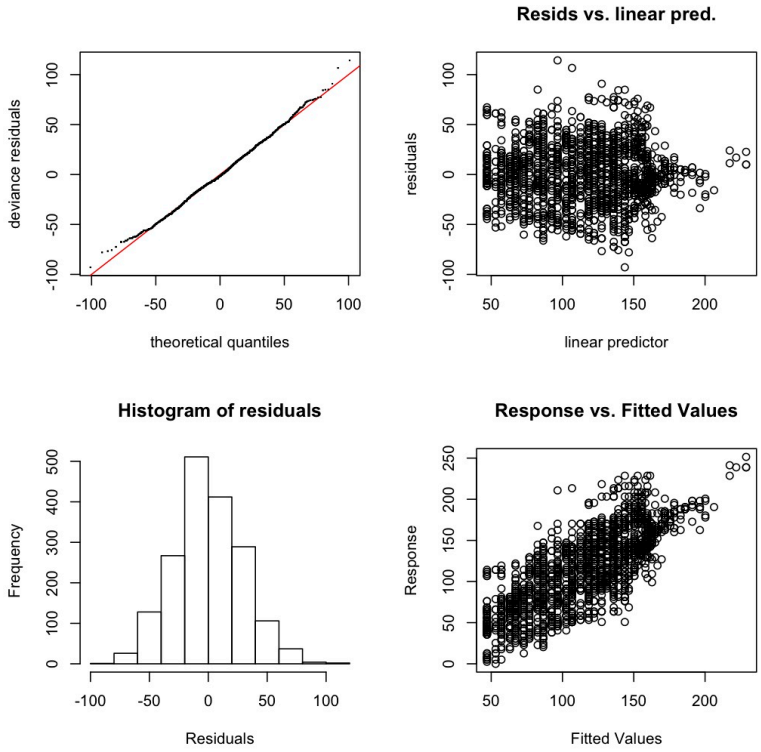



Figure A 7. Residual plots for a model with a smoothing function for snow depth.

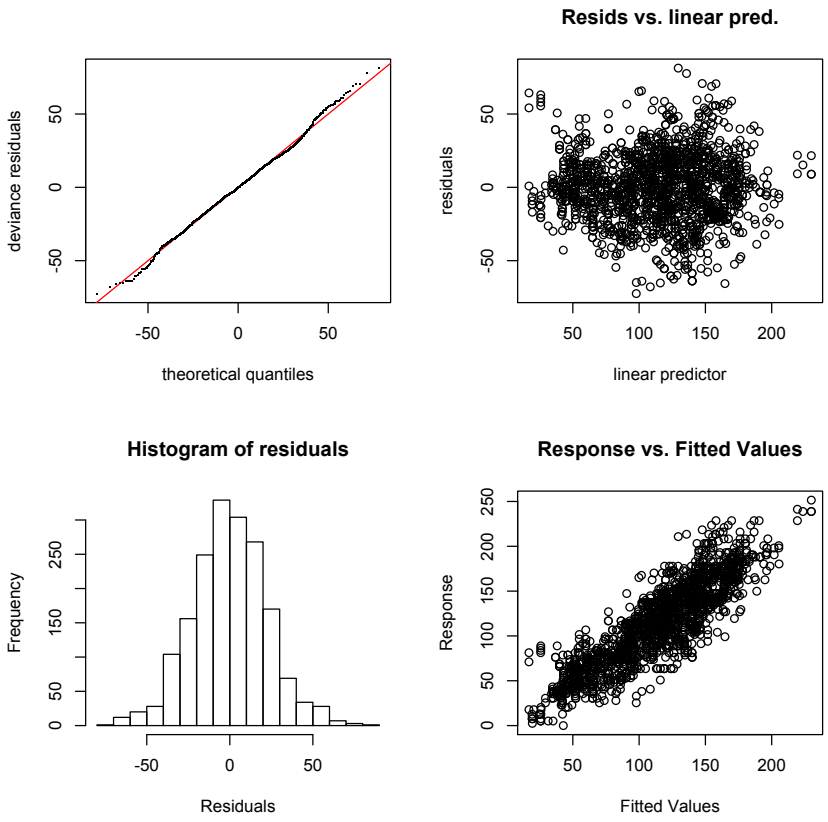



Figure A 8. Residuals from final model for snow depth.

# APPENDIX B

## SNOWPIT PROFILES FROM THE 2011-2012 SEASON

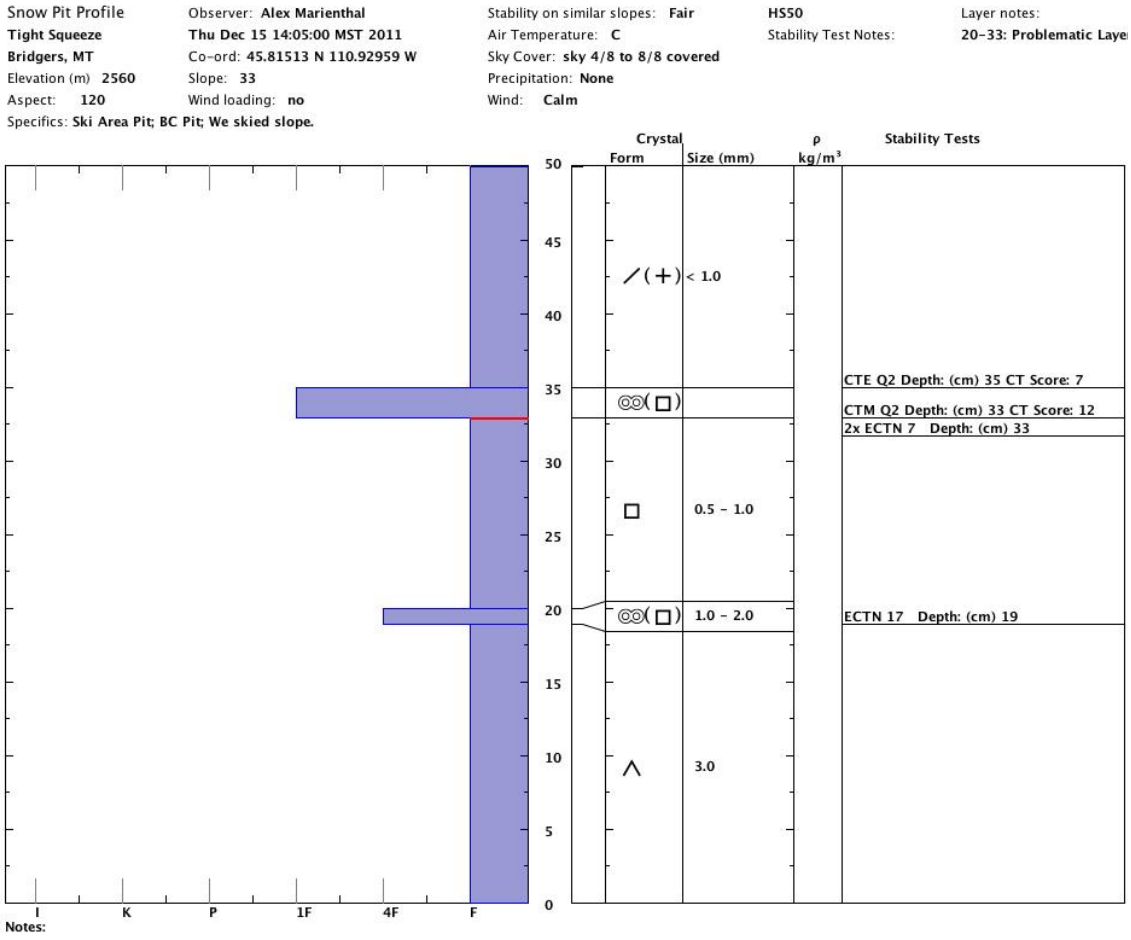



Figure A 23. Snowpit profile from December 15, 2011 at site A.

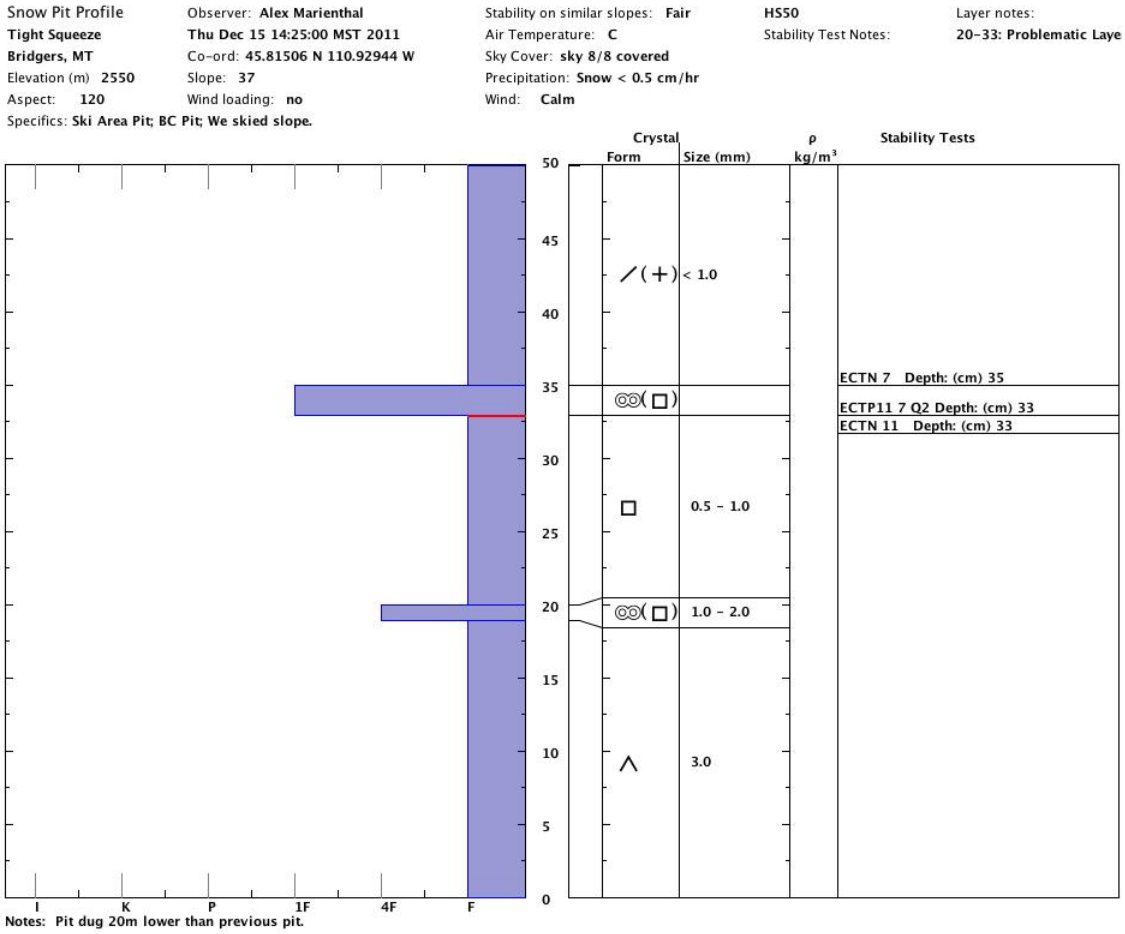



Figure A 24. Second Snowpit profile from December 15, 2011 at site A.

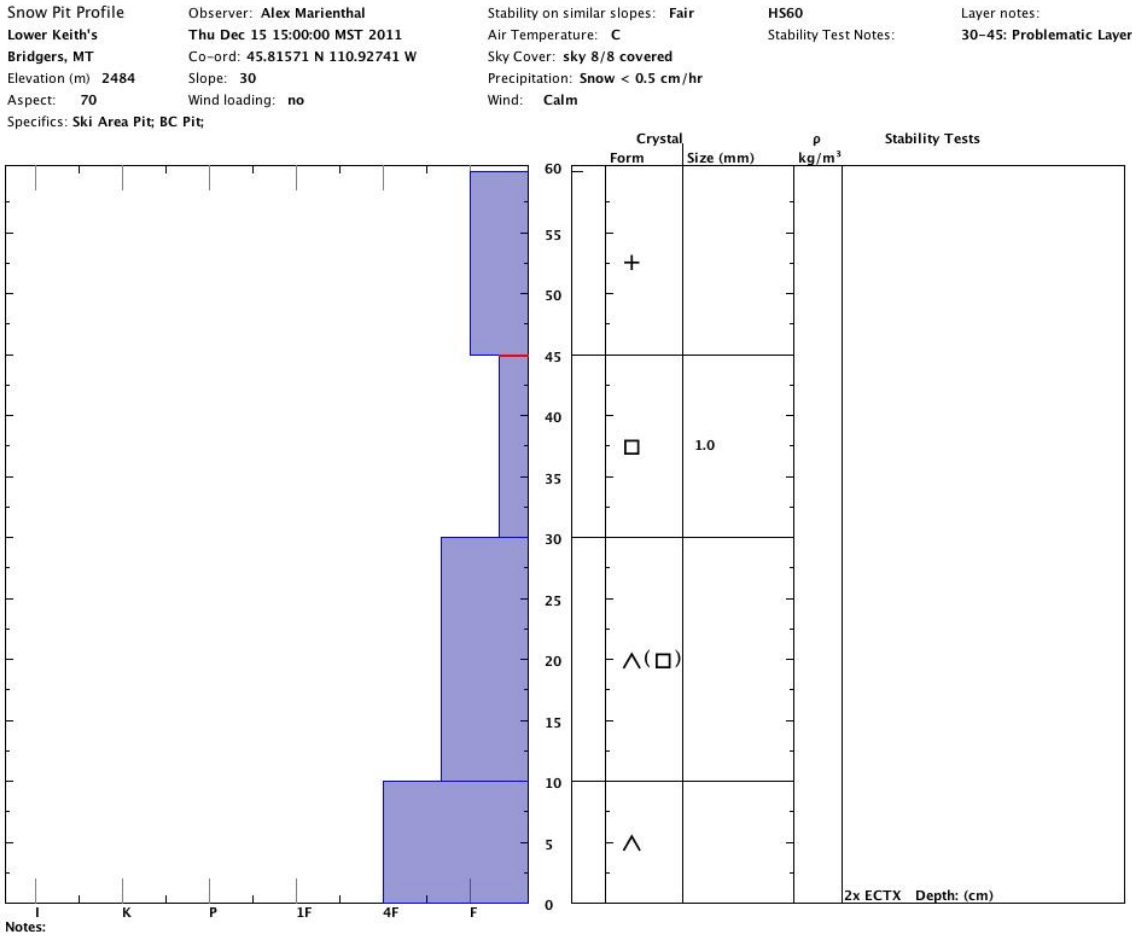



Figure A 25. Snowpit profile on December 15, 2011 north of site B (Lower Keith's starting zone).

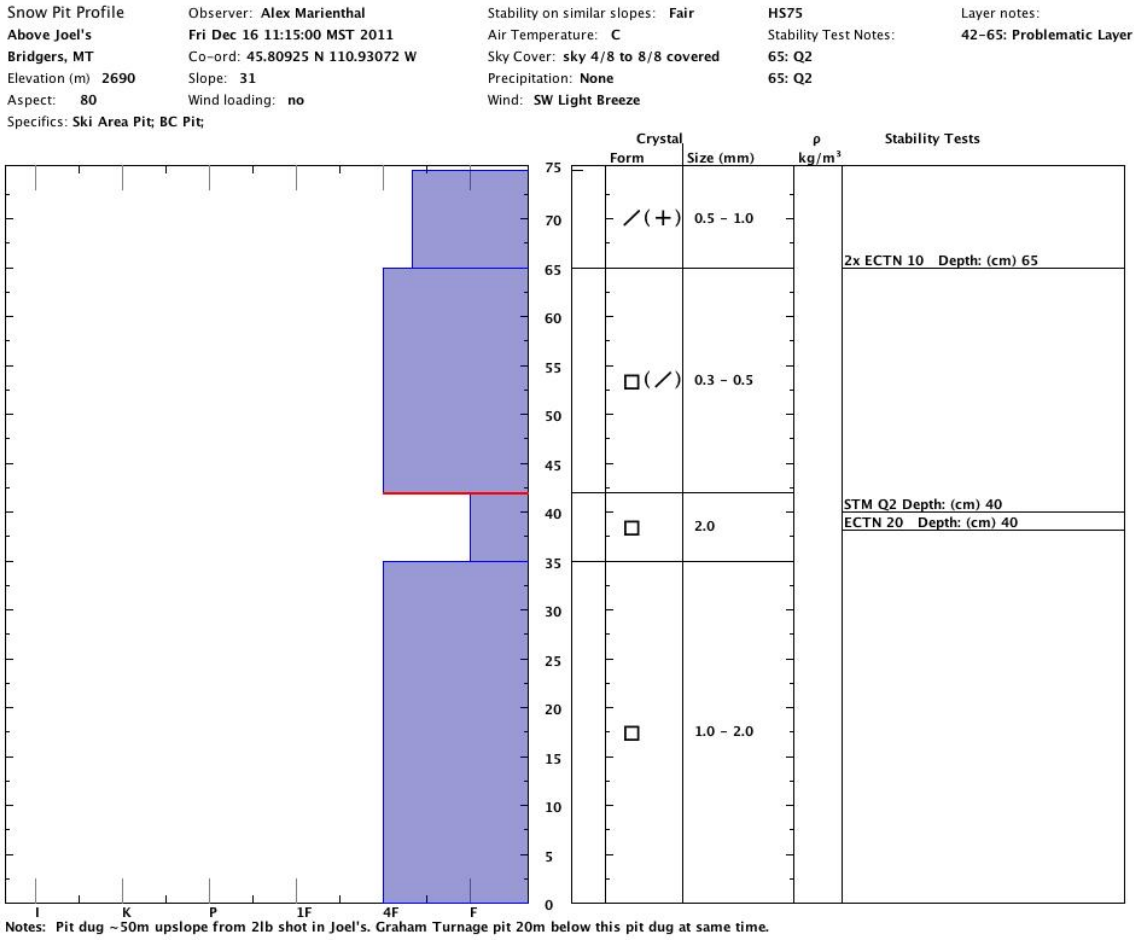



Figure A 26. Snowpit profile on December, 16th, 2011 at Joel's, not shown in Figure A 12 or Table 12.

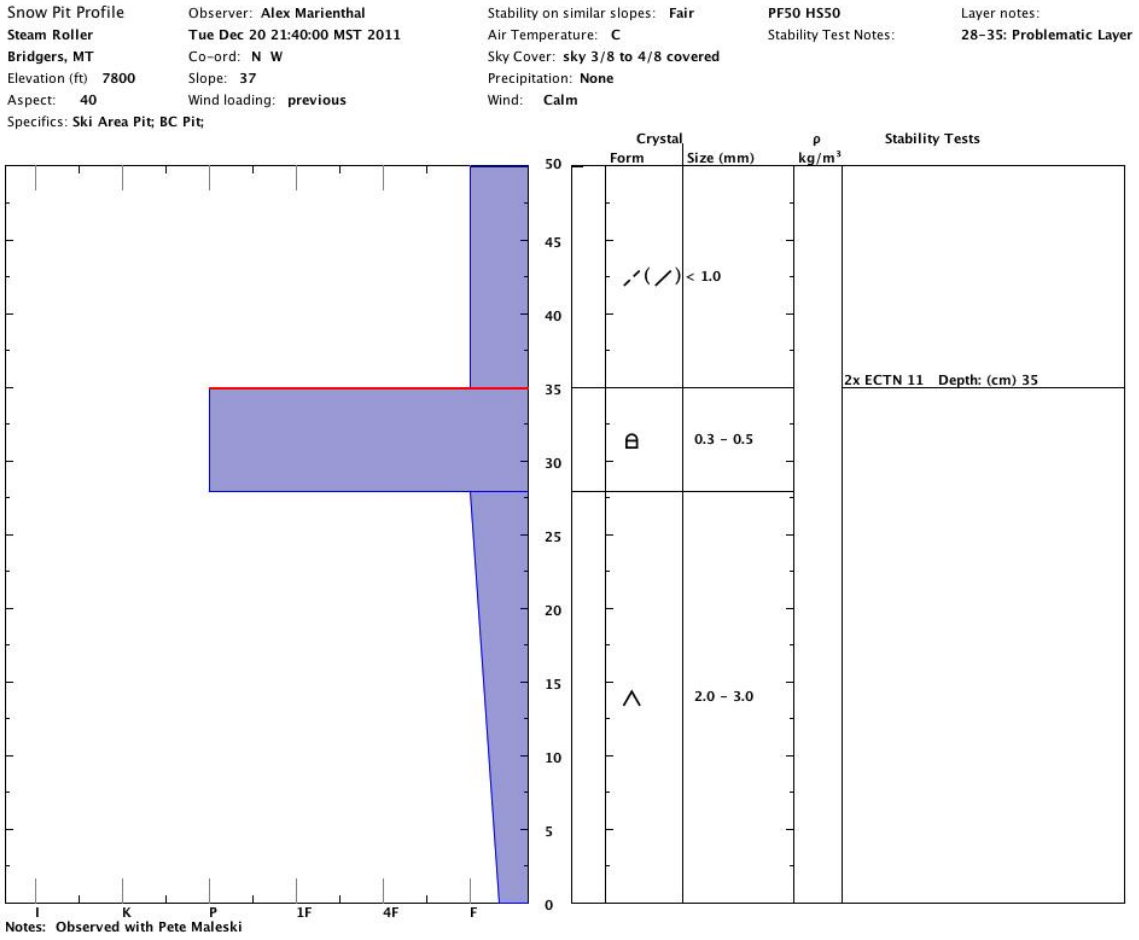



Figure A 27. Snowpit profile at site G on December 20, 2011.

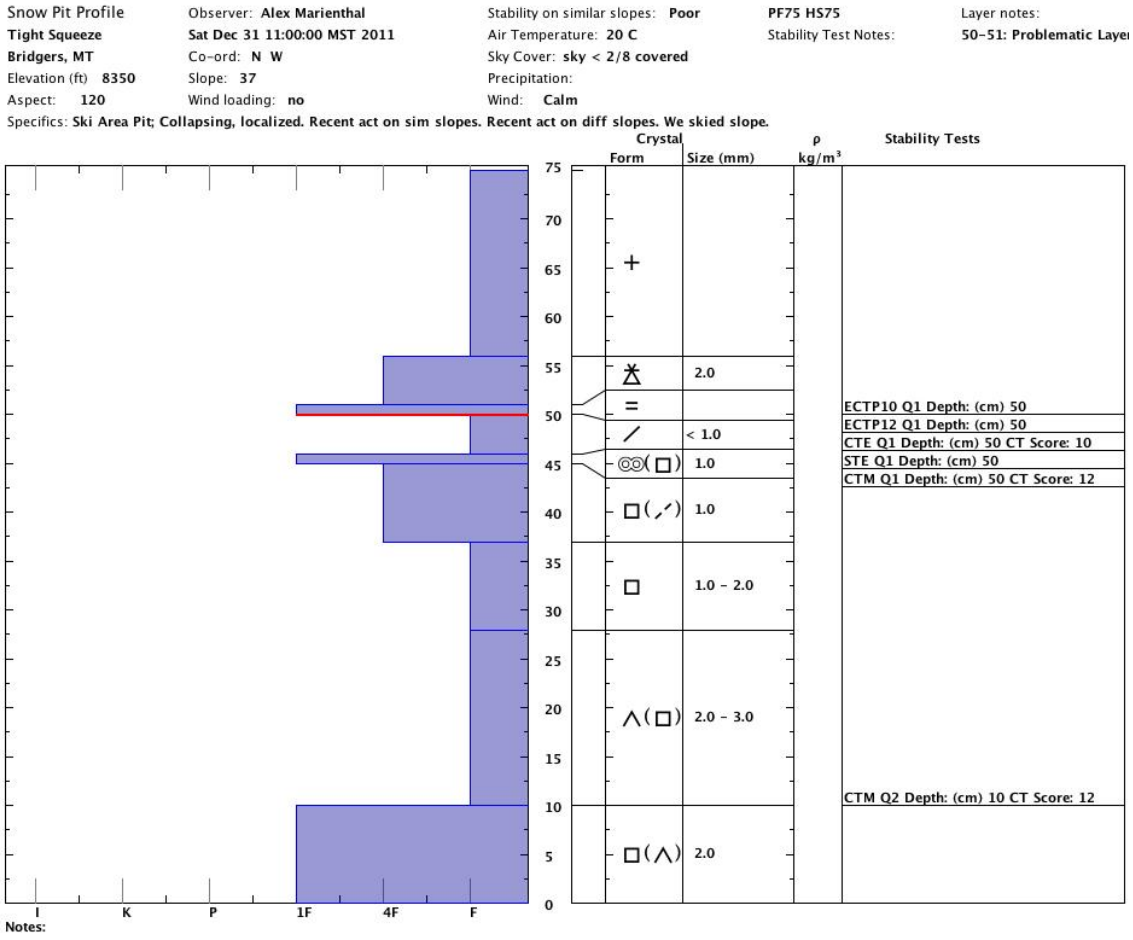



Figure A 28. Snowpit profile from site A on December 31, 2011.

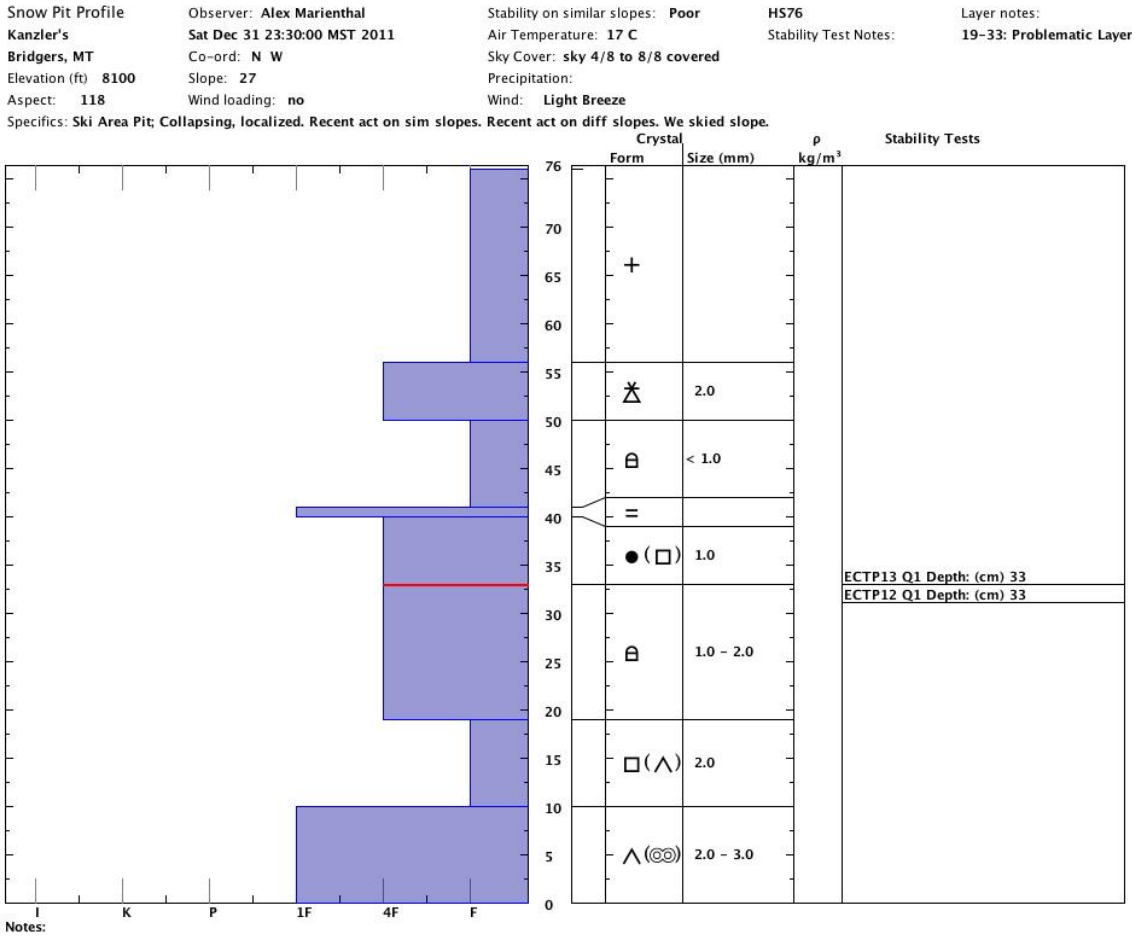



Figure A 29. Snowpit profile from December 31, 2011 at site B.

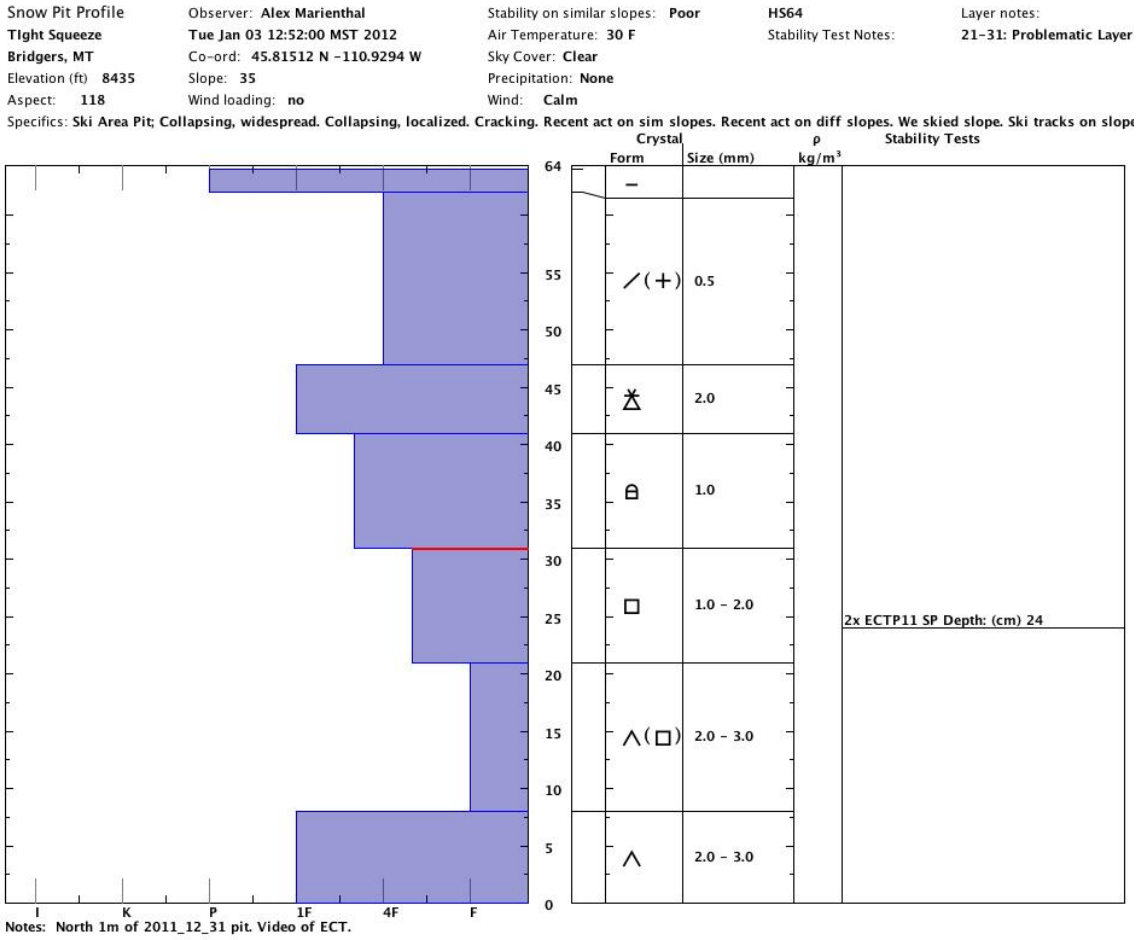



Figure A 30. Snowpit profile from site A on January 3rd, 2012.

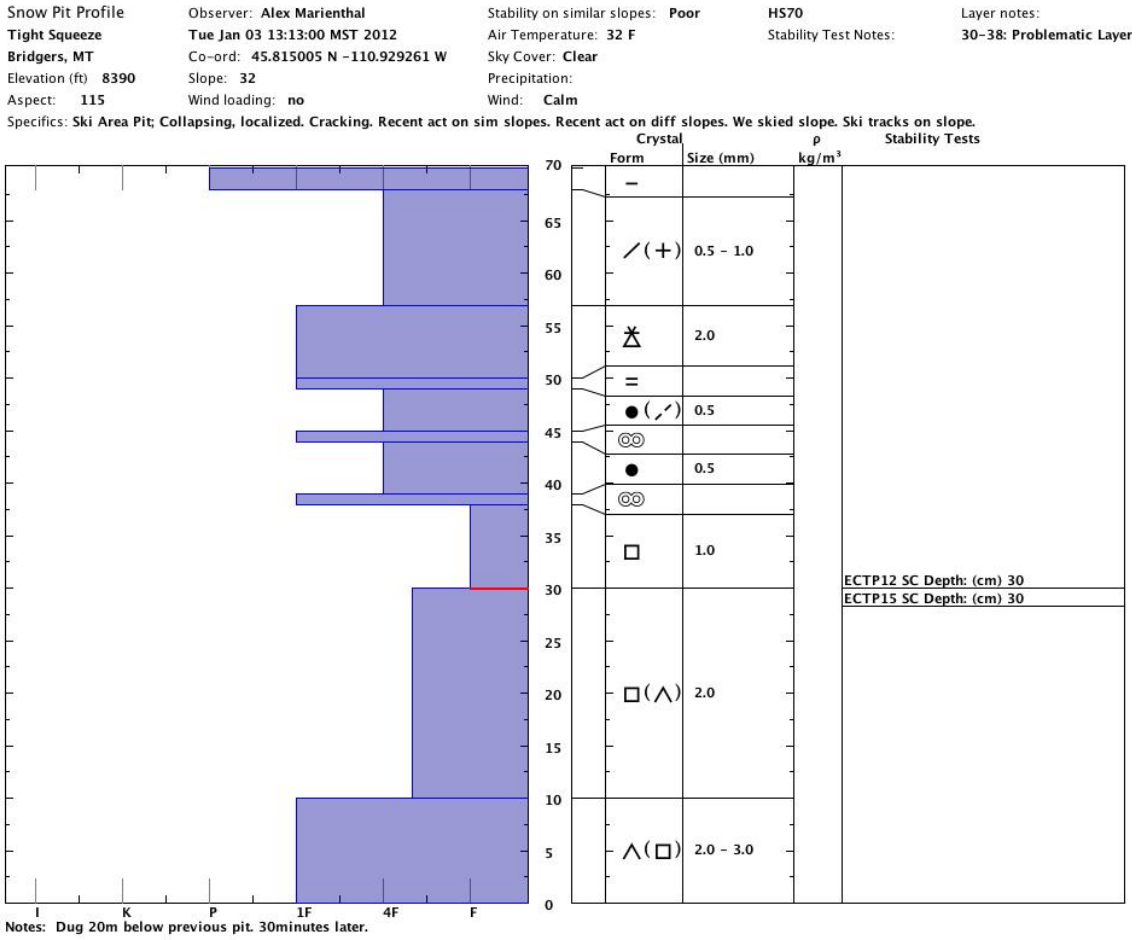



Figure A 31. Second pit from site A on January 3rd, 2012.

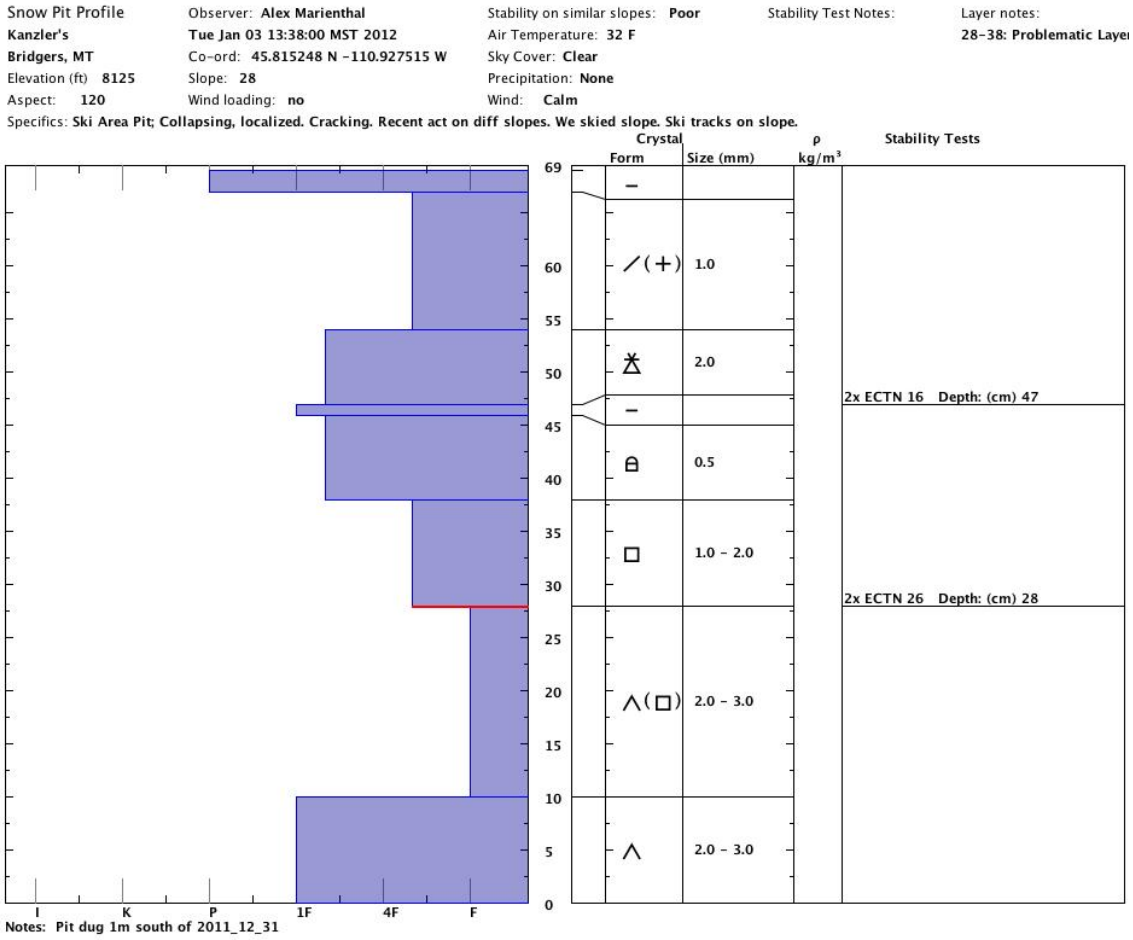



Figure A 32. Snowpit profile on January 3rd, 2012 at site B.

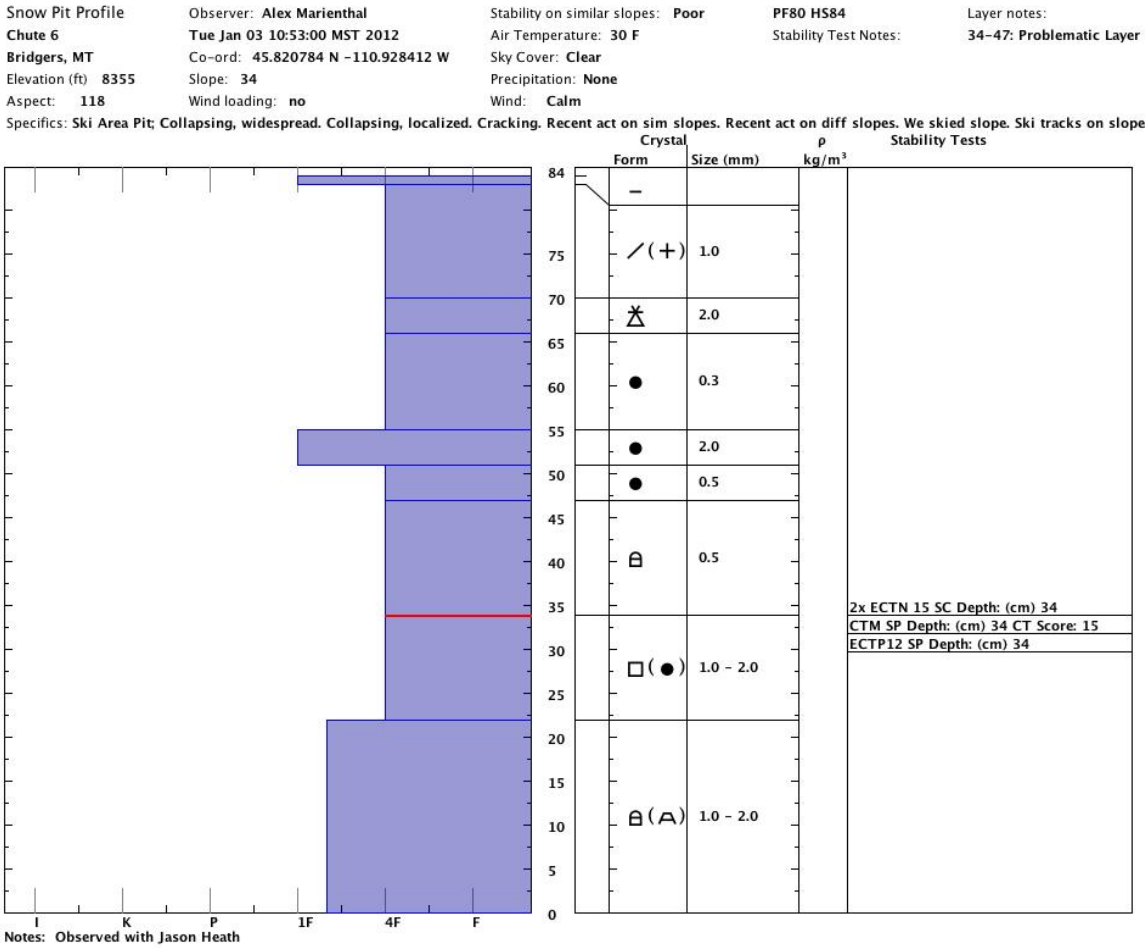



Figure A 33. Snowpit profile from January 3rd, 2012 at site C.

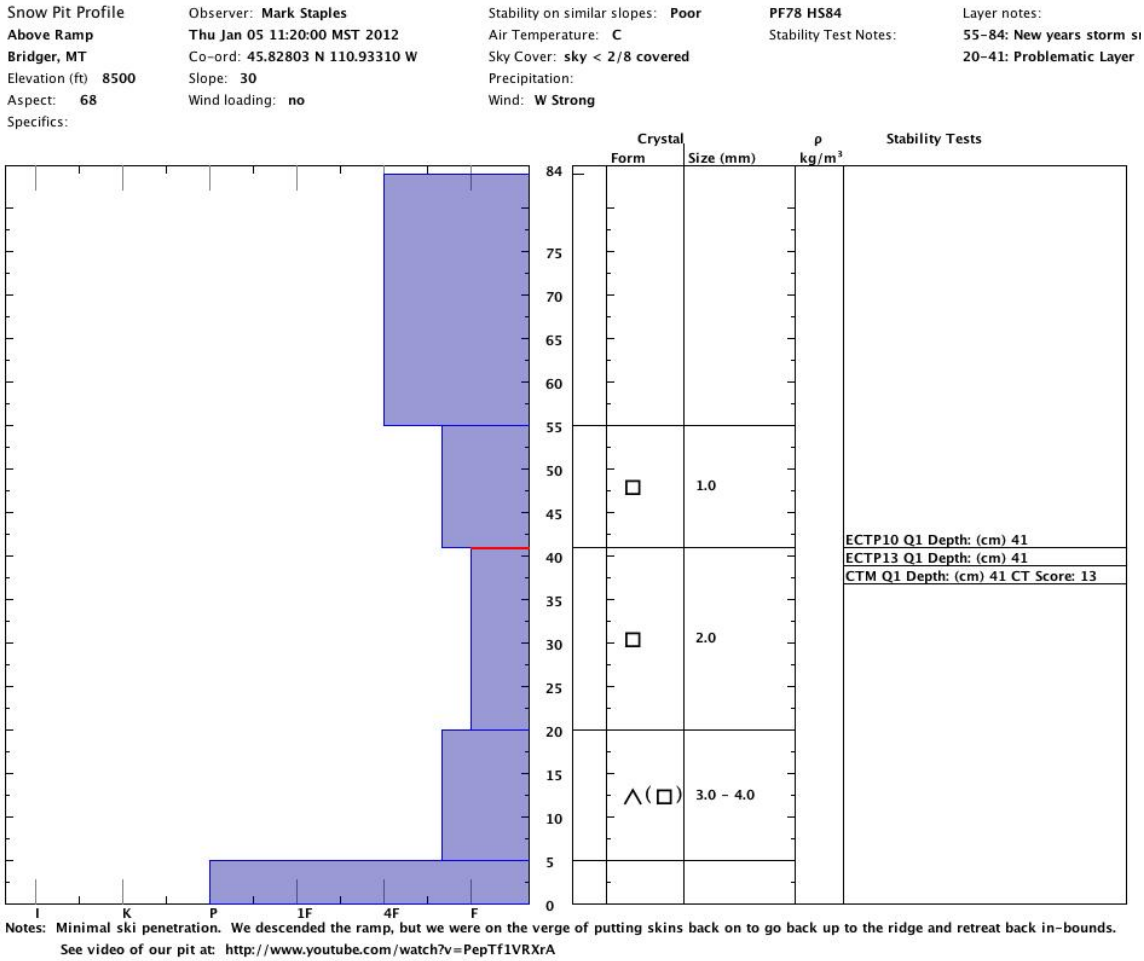



Figure A 34. Snowpit profile on January 5th, 2012 at site D.

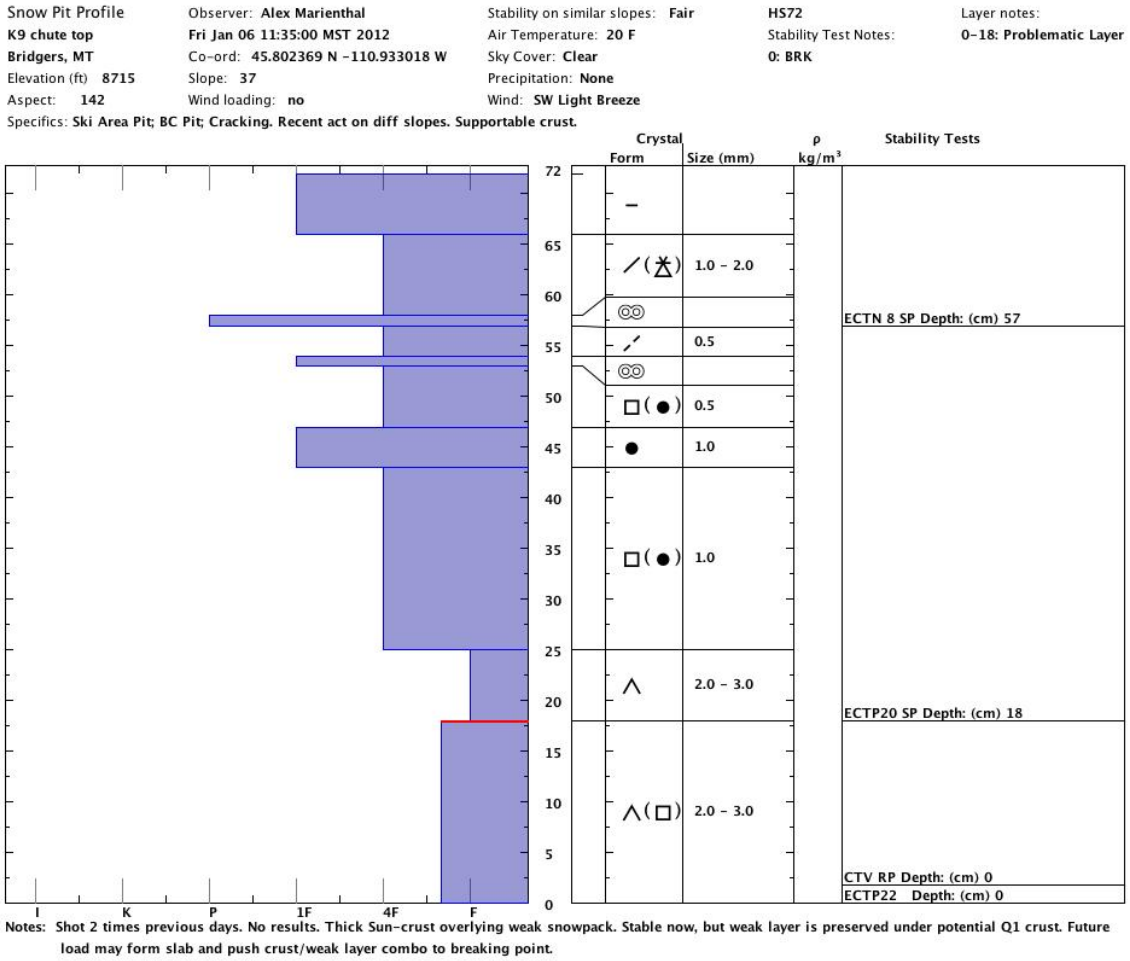



Figure A 35. Snowpit profile from January 6th, 2012 at site E.

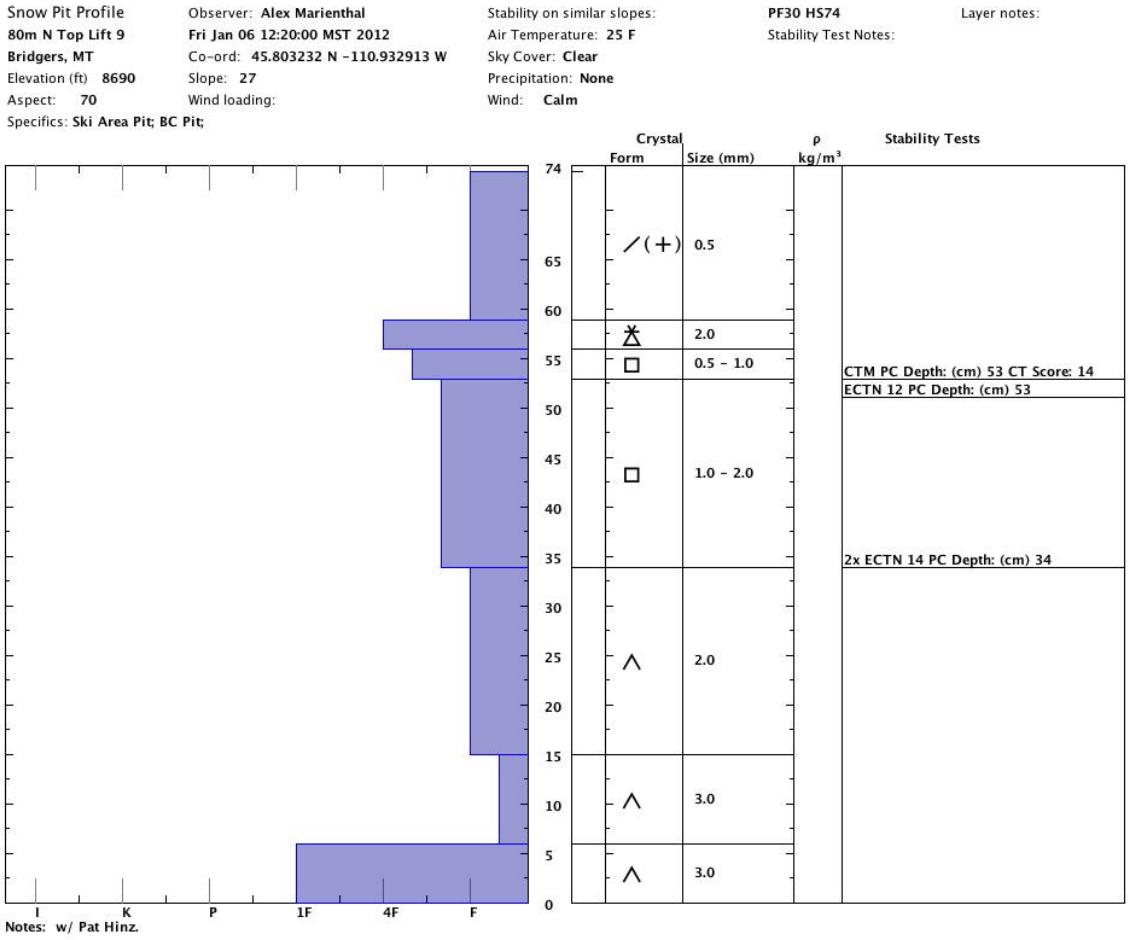



Figure A 36. Second pit from site E on January 6th, 2012.

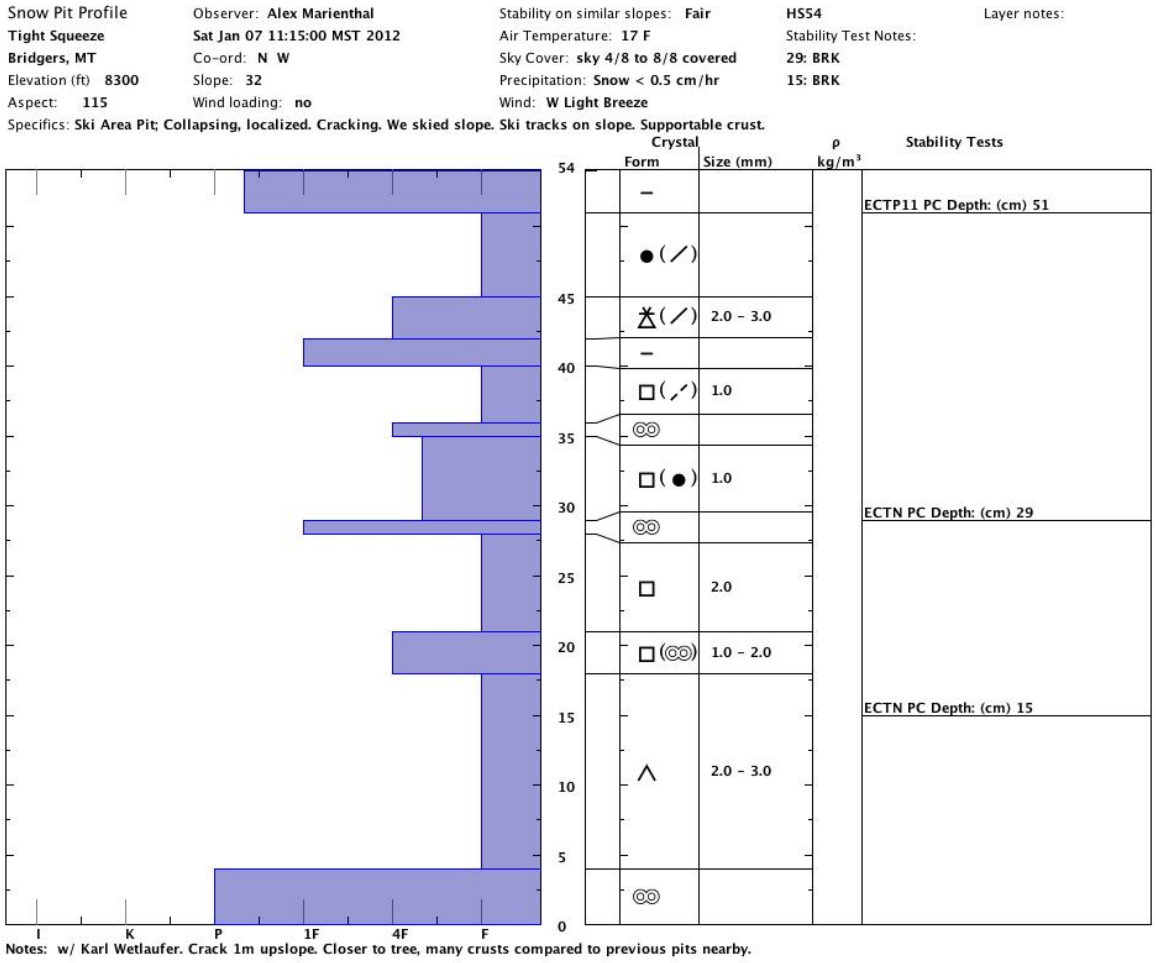



Figure A 37. Snowpit profile from site A on January 7th, 2012.

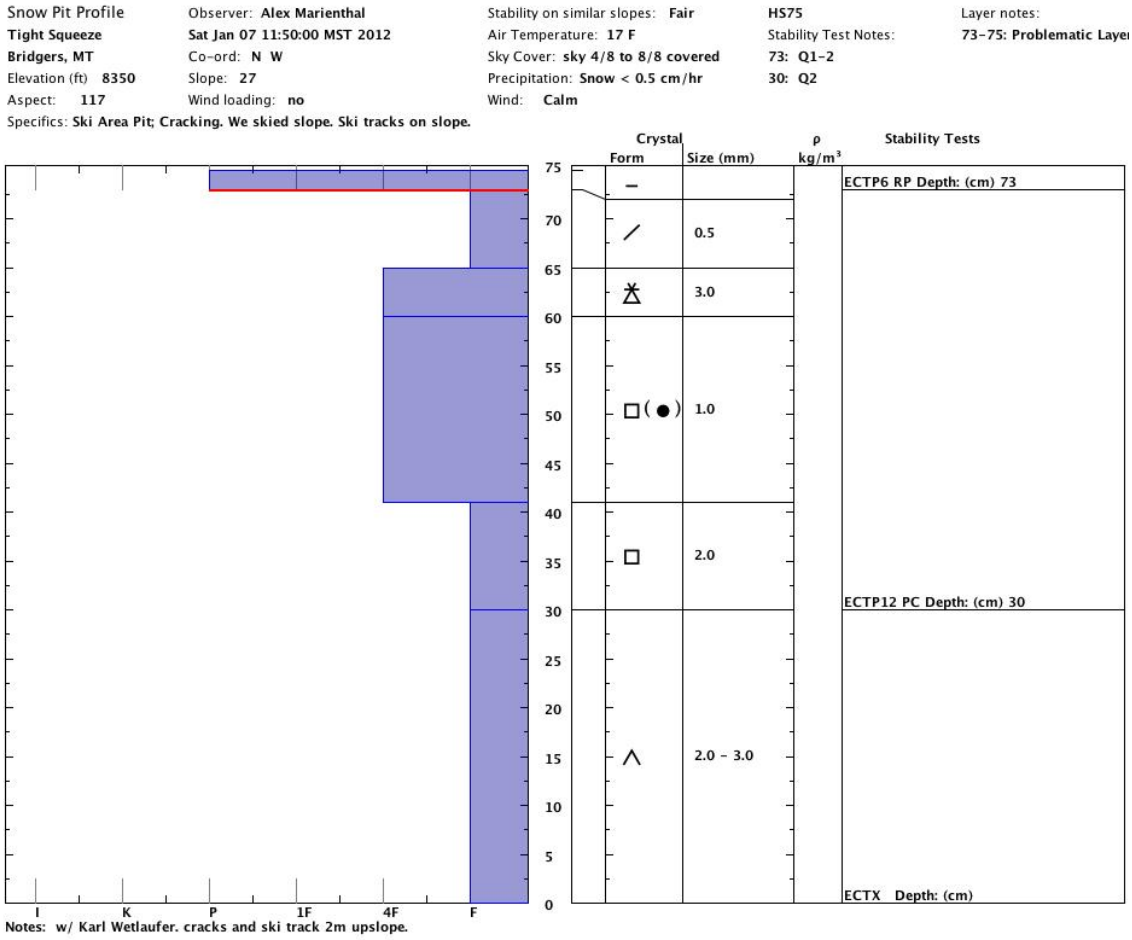



Figure A 38. Second pit from site A on January 7th, 2012.

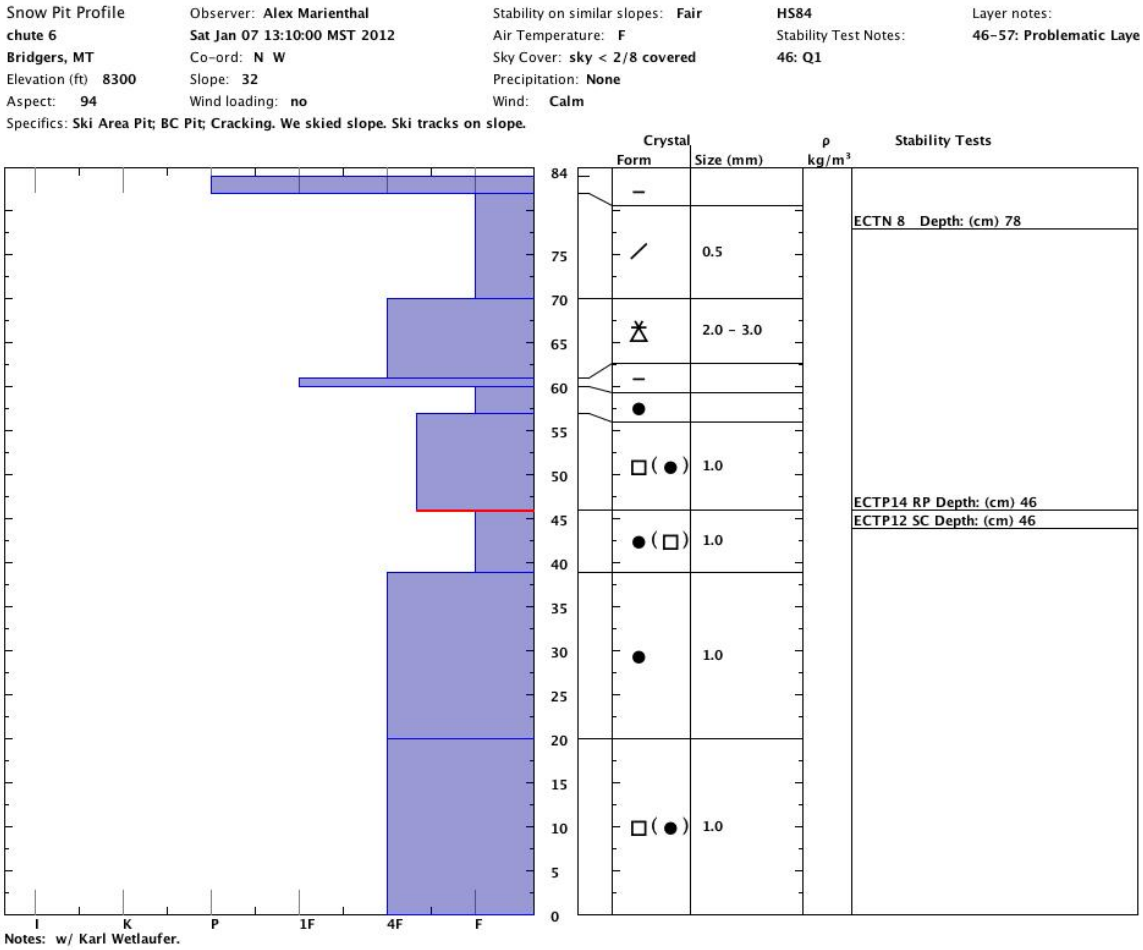



Figure A 39. Snowpit profile from site C on January 7, 2012.

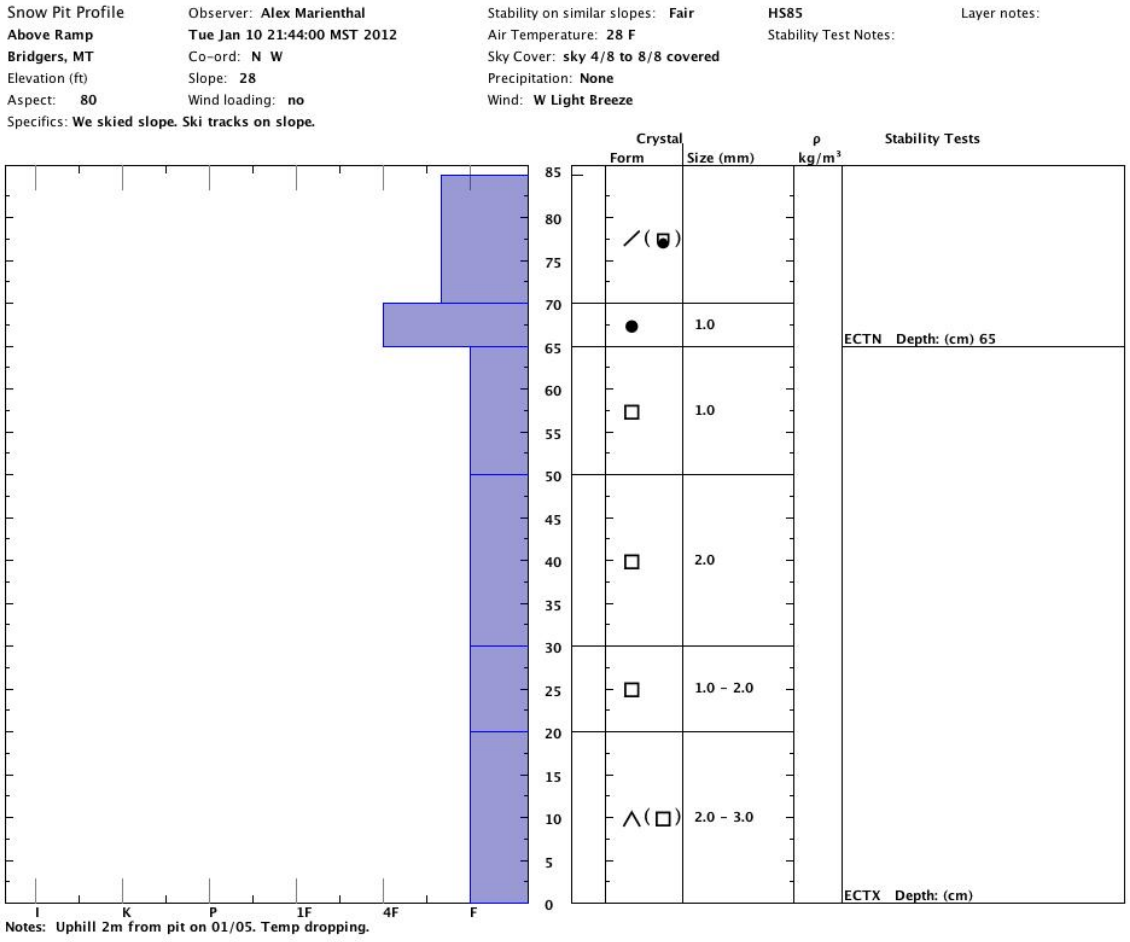



Figure A 40. Snowpit profile from site D on January 10th, 2012.

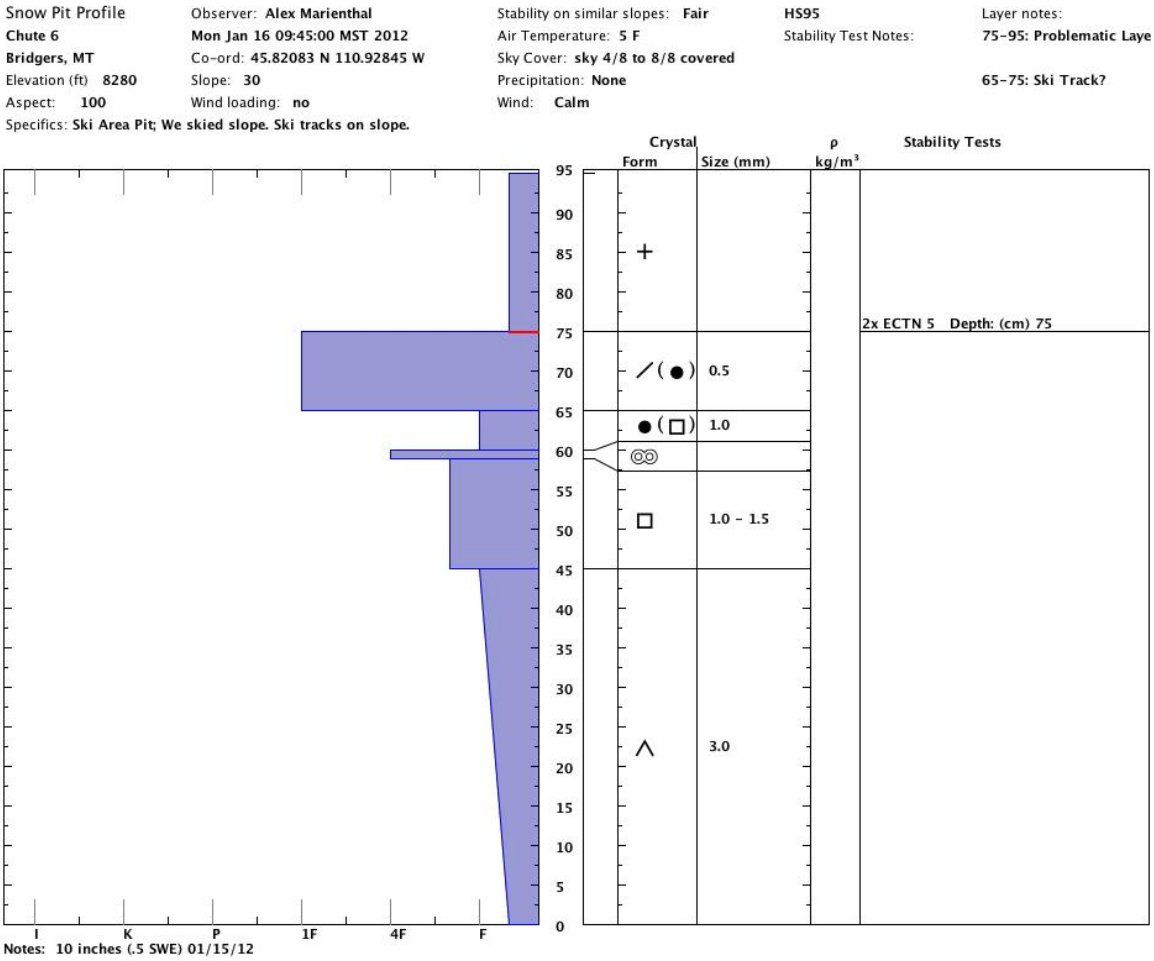



Figure A 41. Snowpit profile at site C on January 16th, 2012.

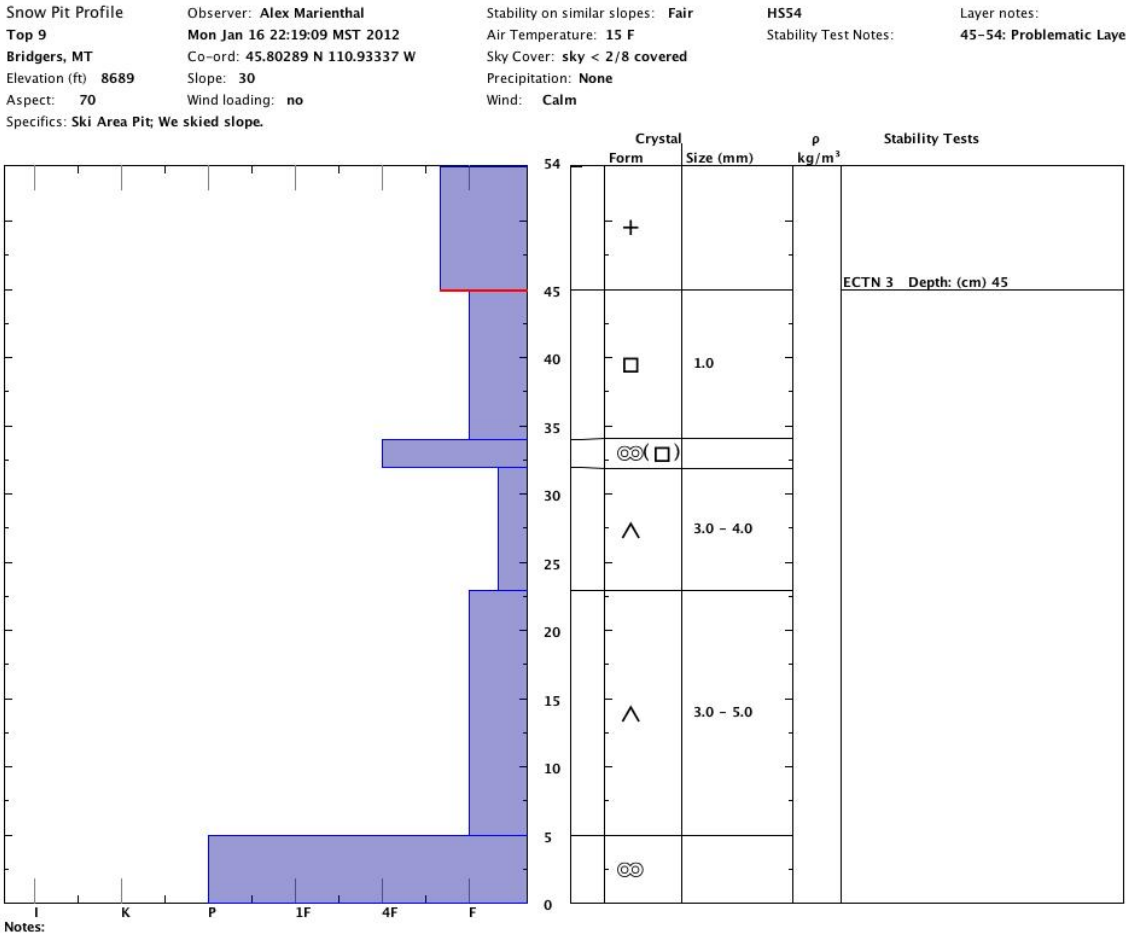



Figure A 42. Snowpit profile from site E on January 16th, 2012.

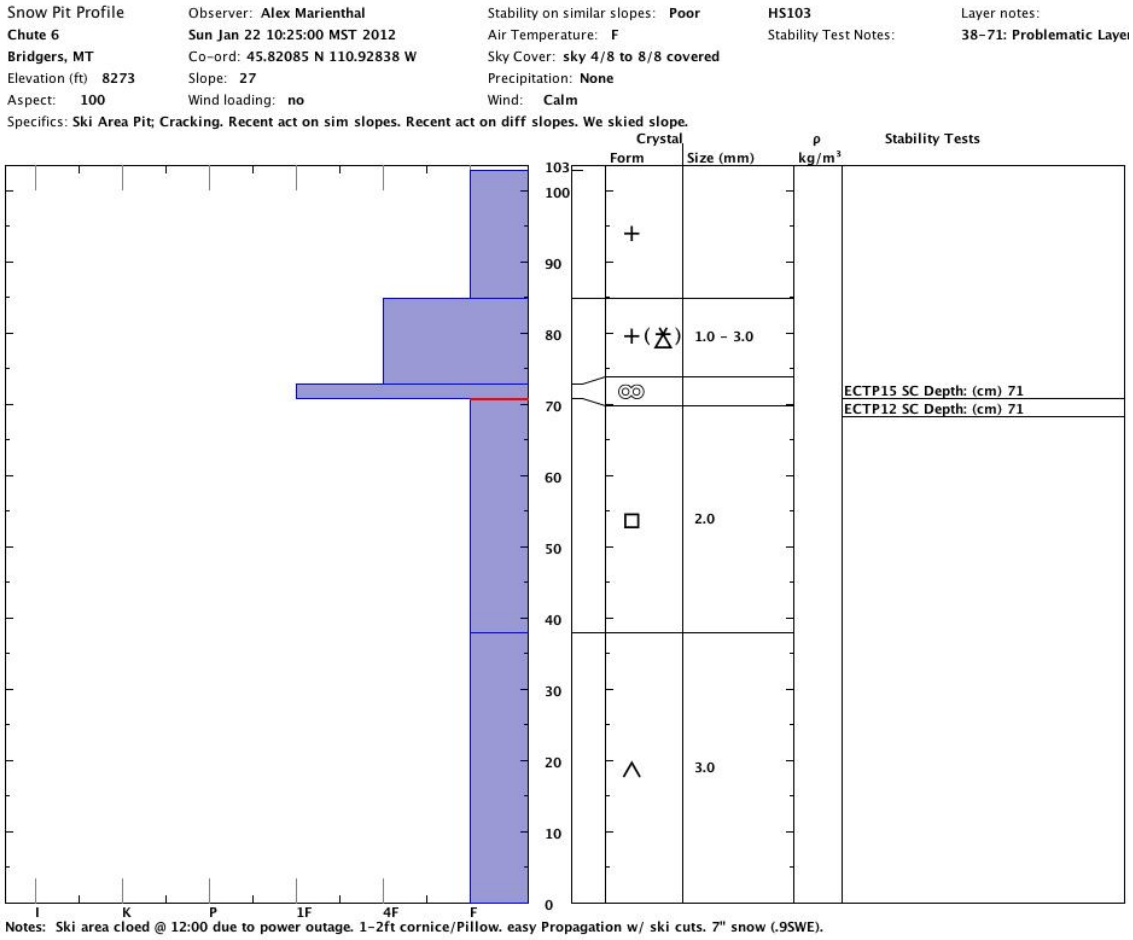



Figure A 43. Snowpit profile from site C on January 22nd, 2012.

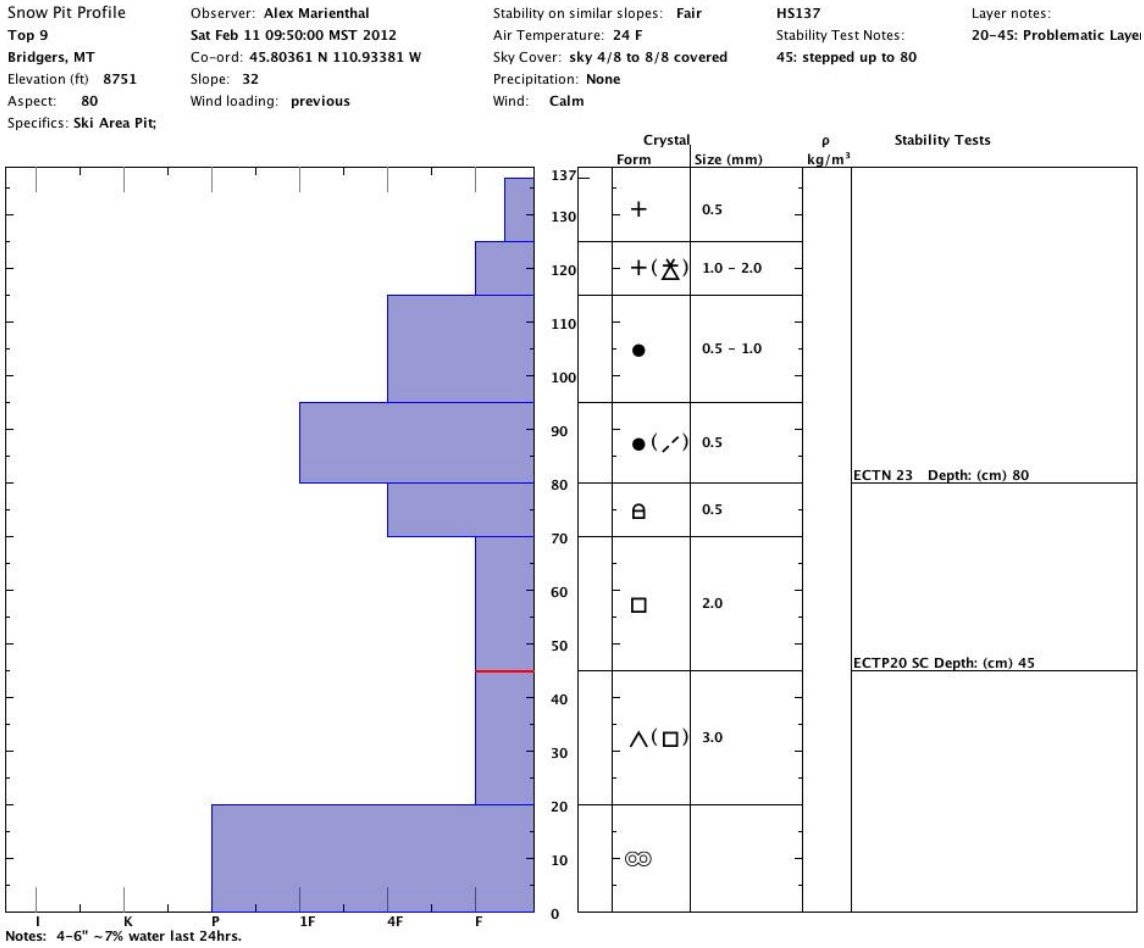



Figure A 44. Snowpit profile from site E on February 11th, 2012.

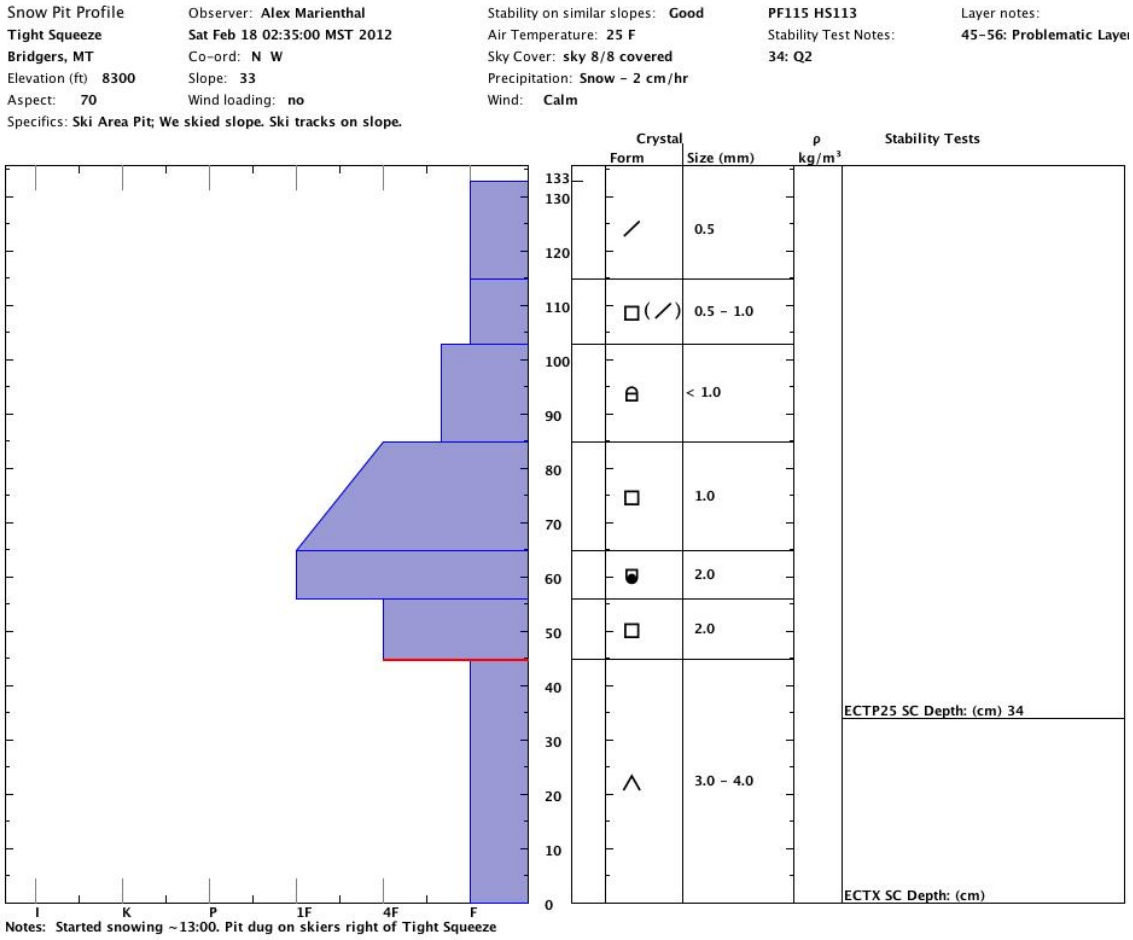



Figure A 45. Snowpit profile at site A on February 18th, 2012.

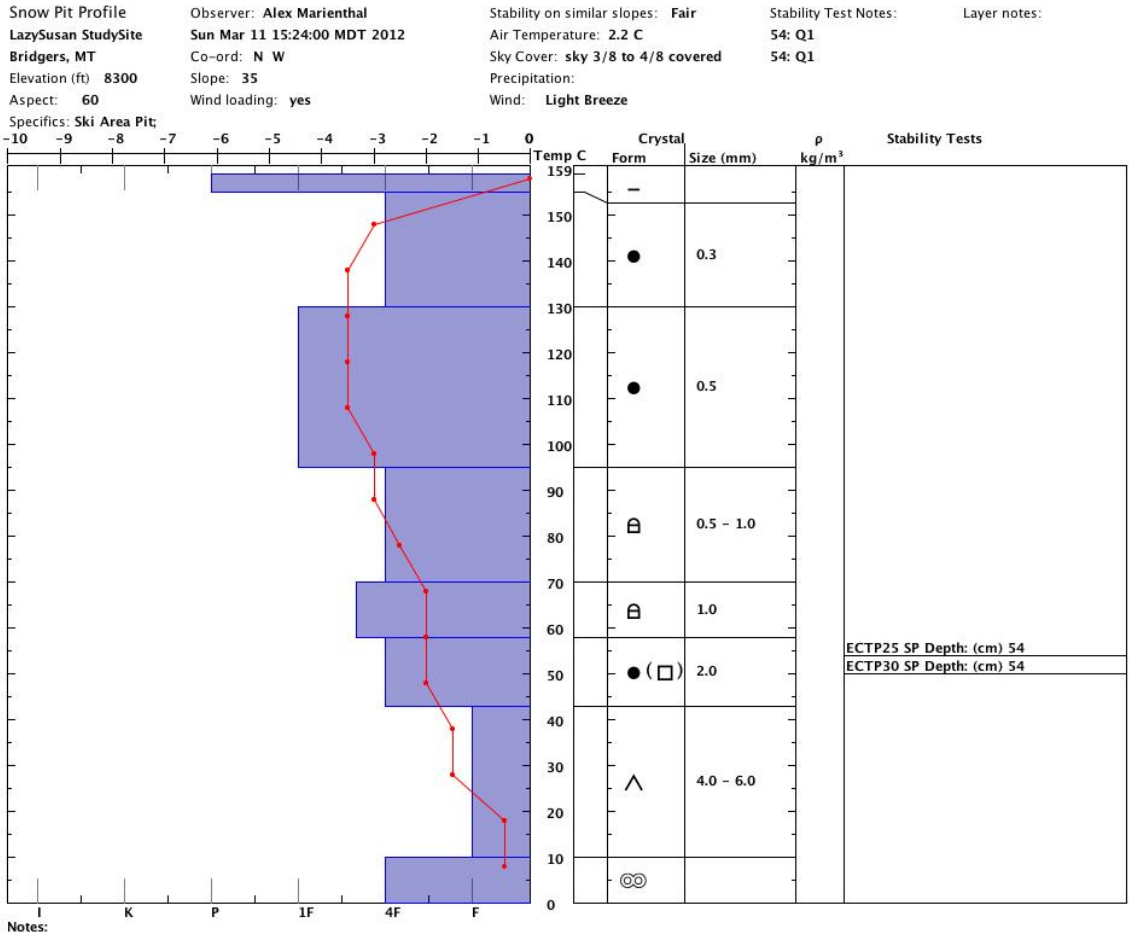



Figure A 46. Snowpit profile at site F on March 11, 2012.

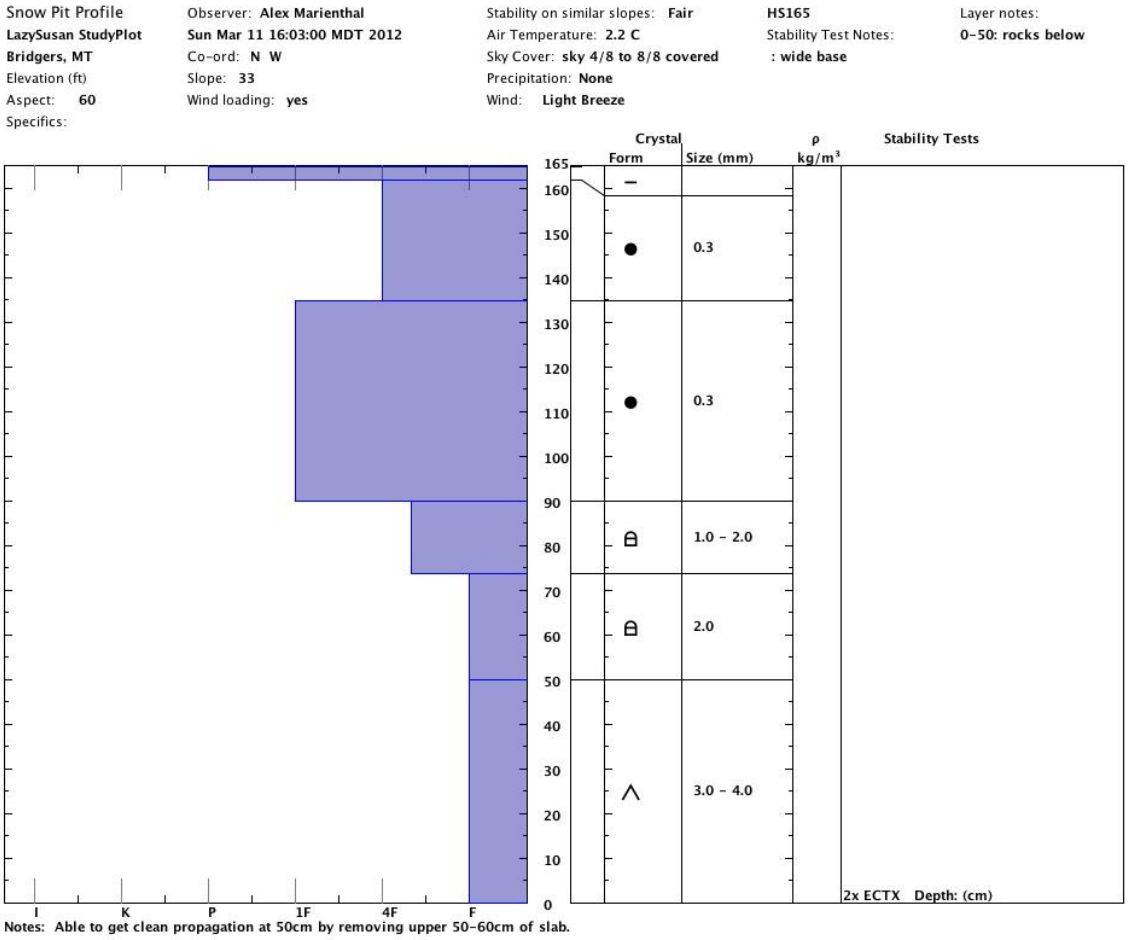



Figure A 47. Second Snowpit profile from site F on March 11th, 2012.

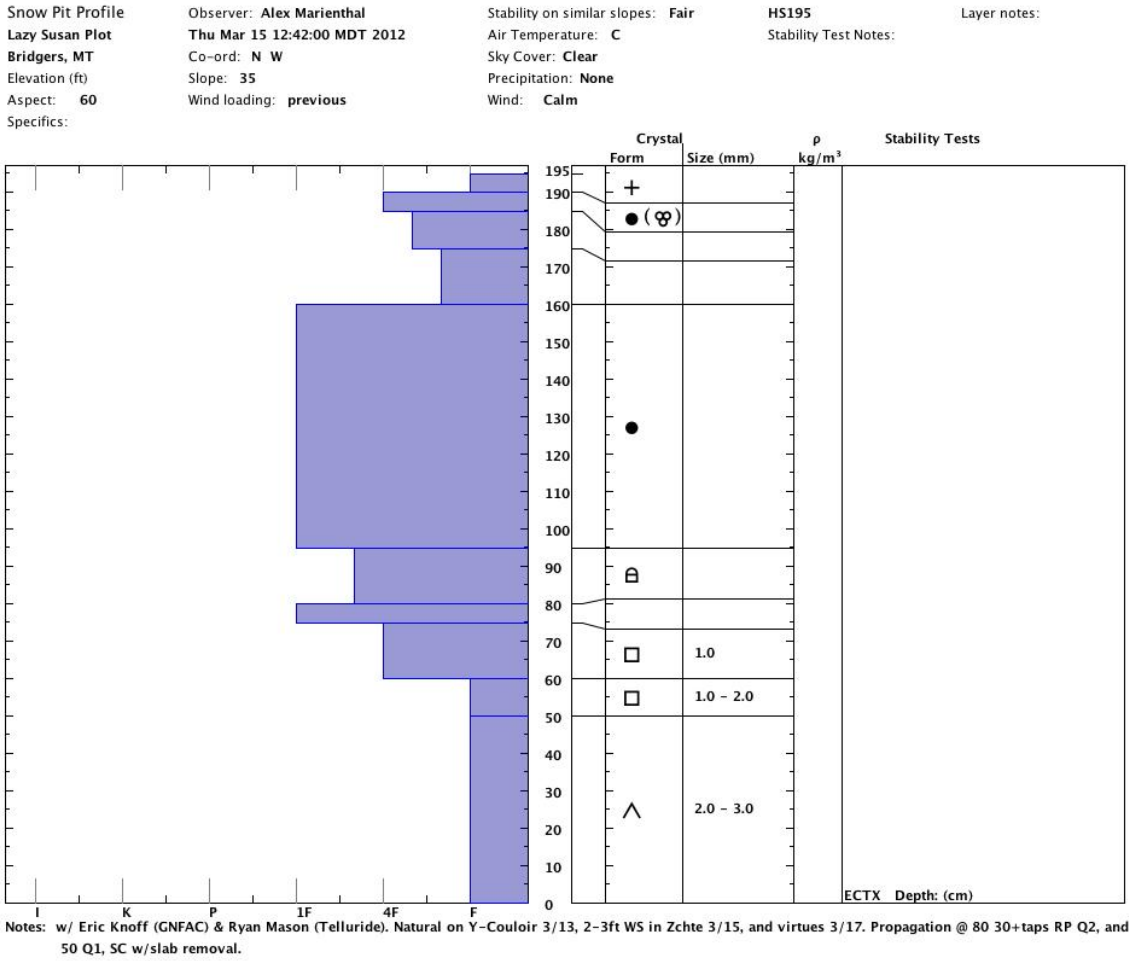



Figure A 48. Snowpit profile from March 15th, 2012 at site F.

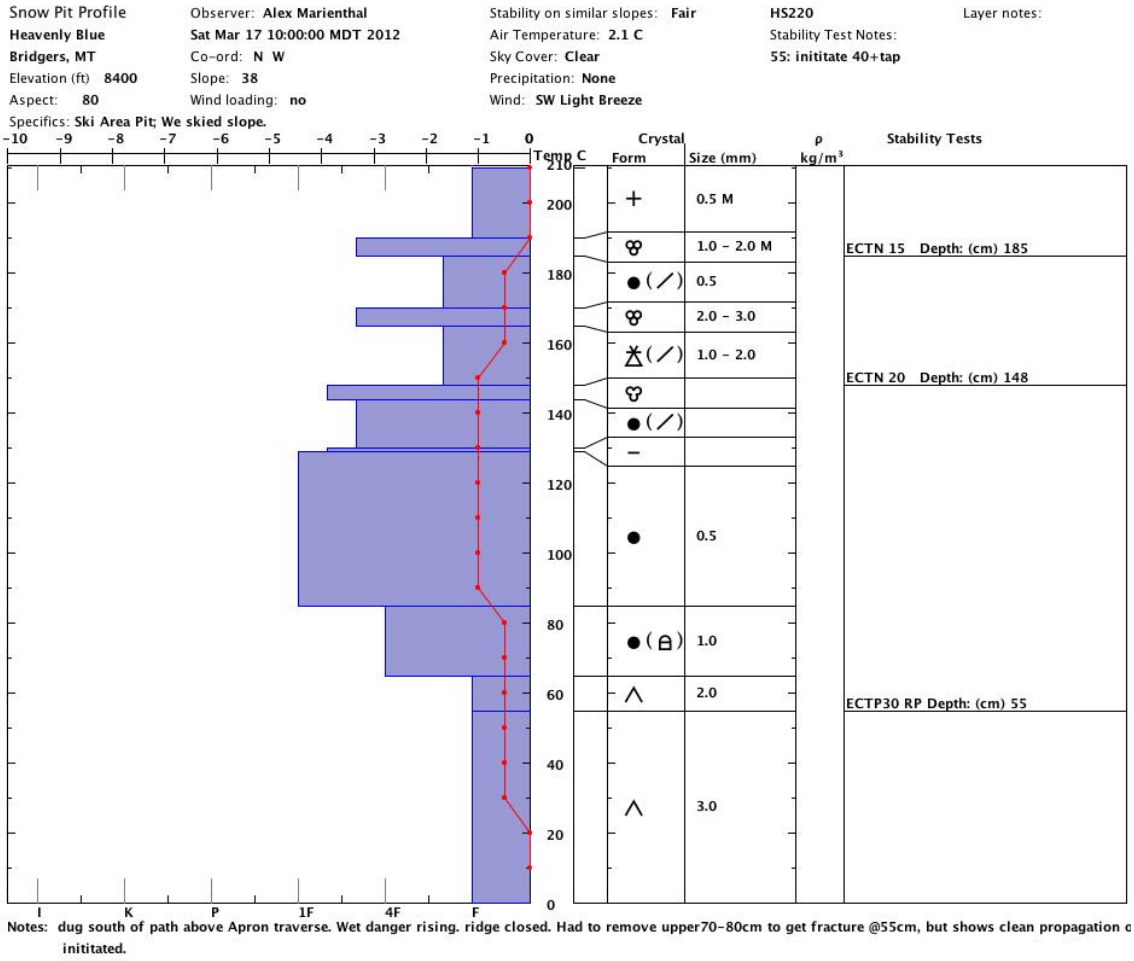



Figure A 49. Snowpit profile from March 17th, 2012 at site G.

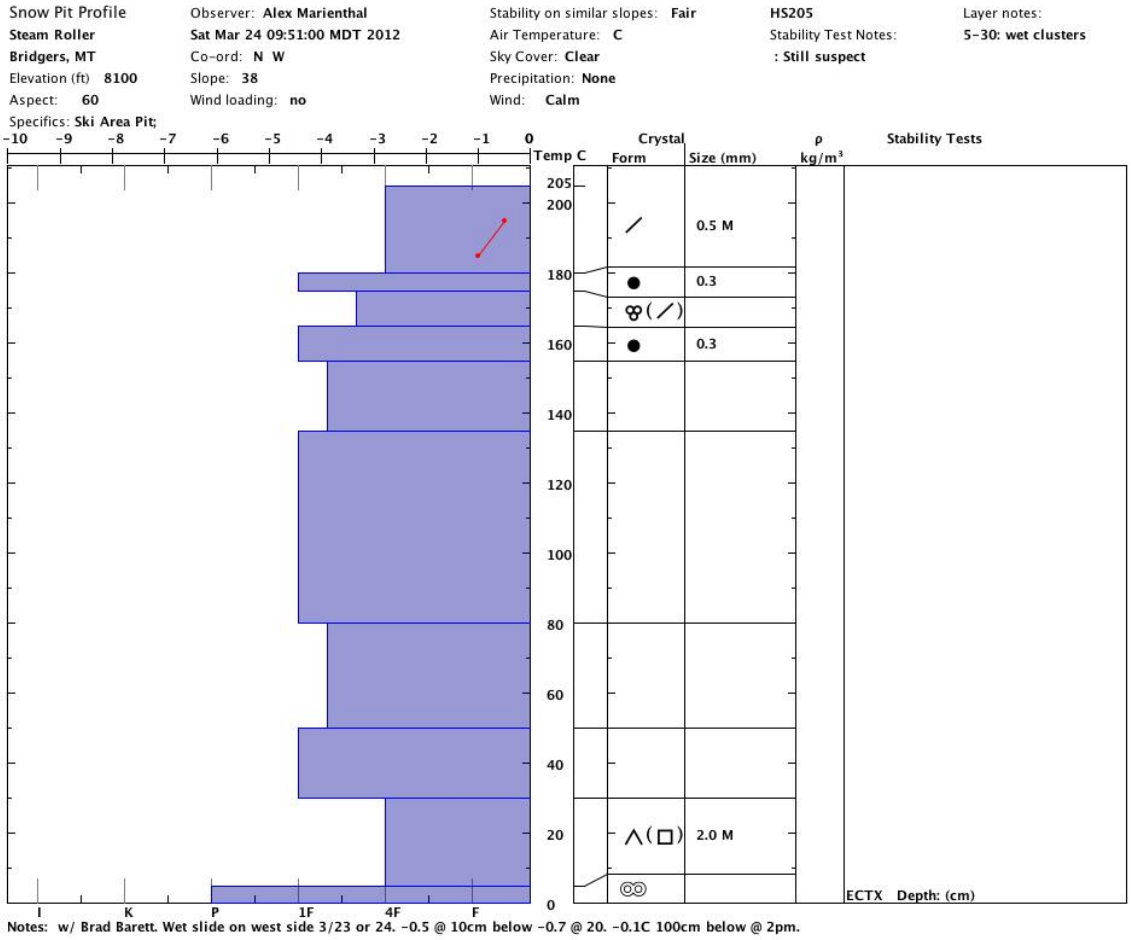



Figure A 50. Snowpit profile from March 24th, 2012 at site G.

# APPENDIX C

## CLASSIFICATION TREES

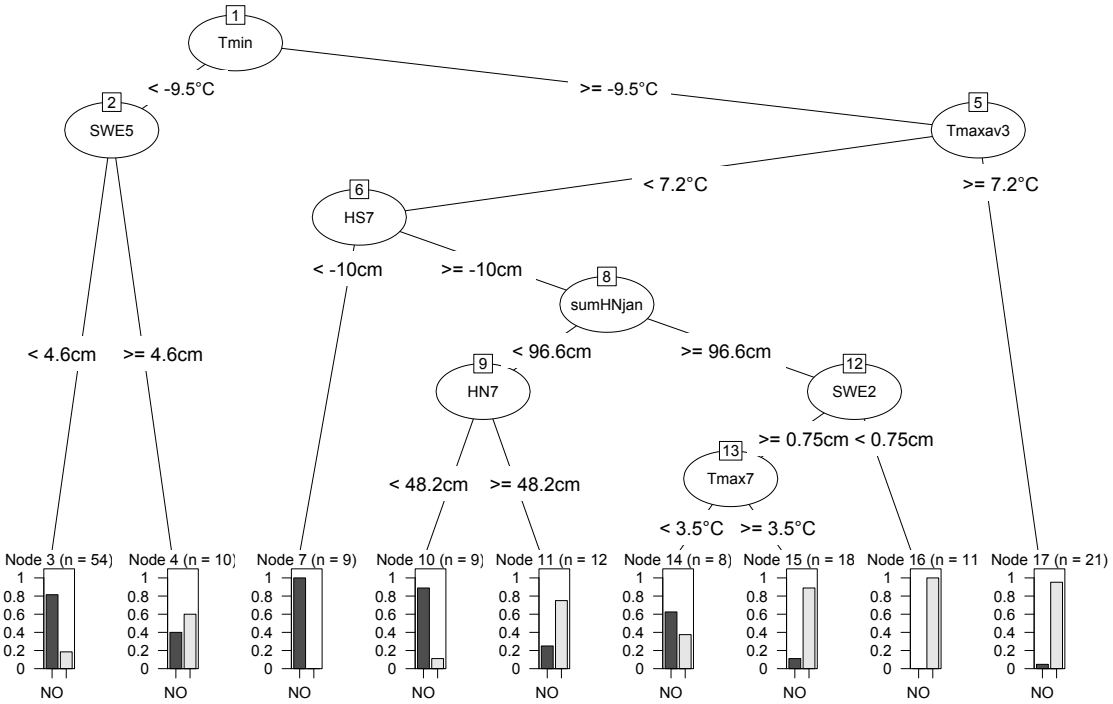



Figure A 51. Over-fit tree using all daily and seasonal metrics. Non-deep slab days are black bars (left), deep slab days are grey bars (right).

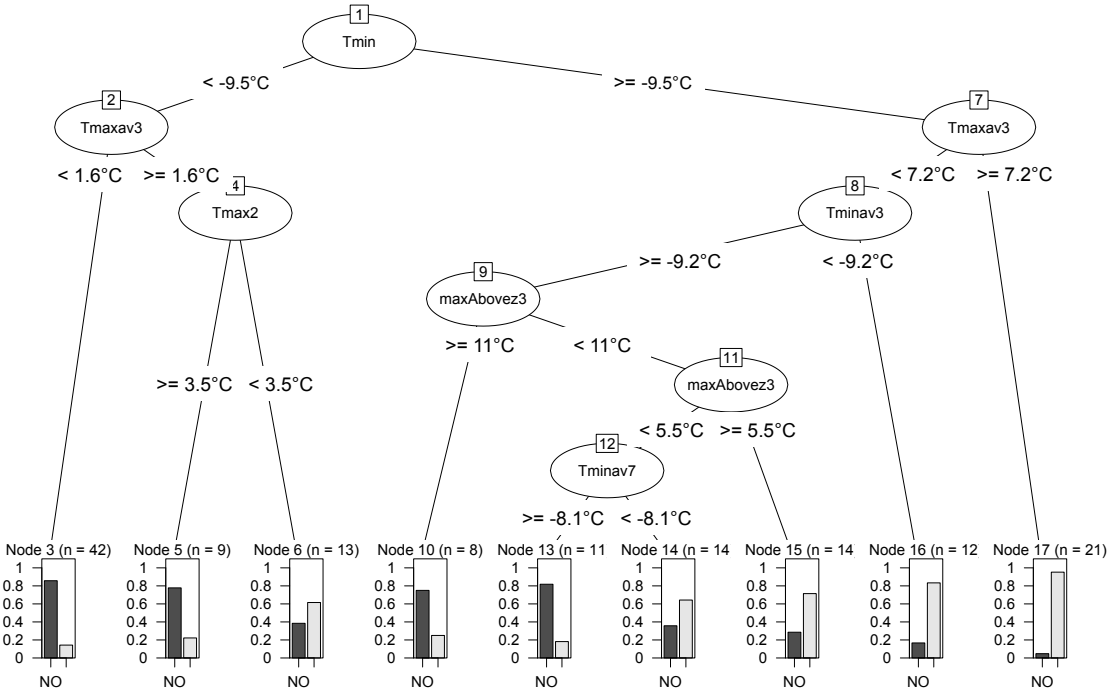



Figure A 52. Over-fit tree using all significant variables. Non-deep slab days are black bars (left), deep slab days are grey bars (right).

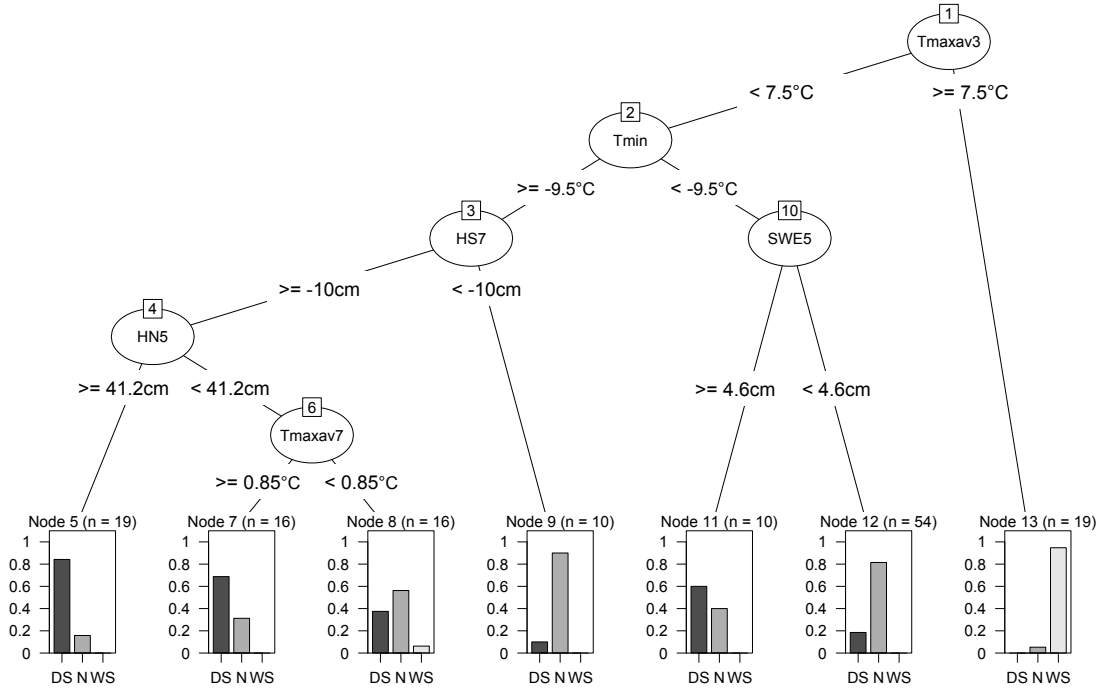



Figure A 53. Over-fit tree created using daily metrics to split days with ether dry or wet deep slabs or no deep slabs.

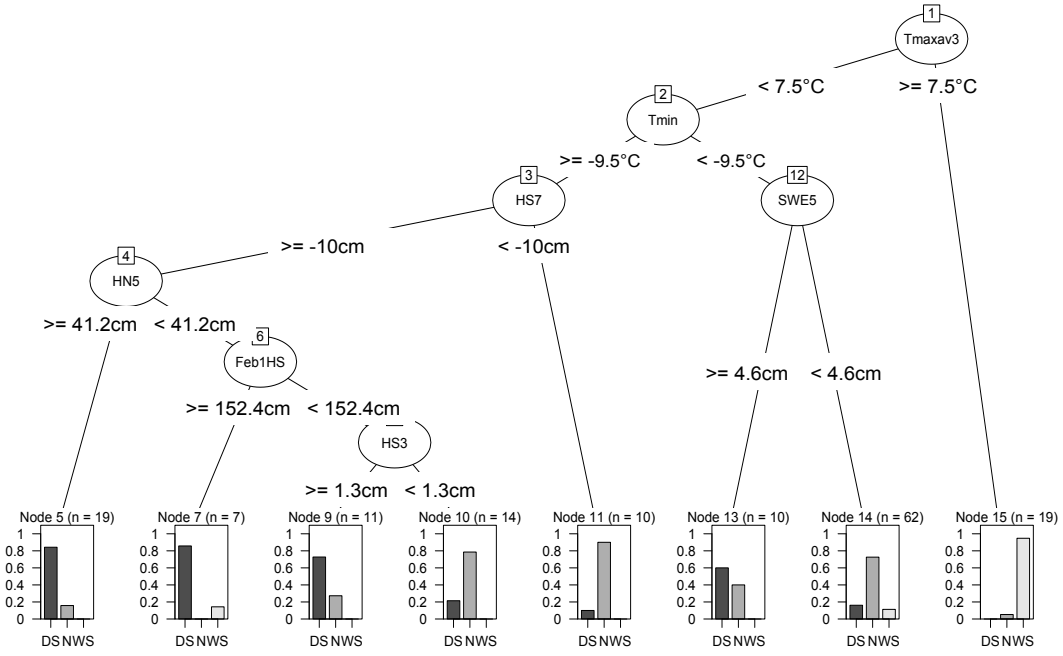



Figure A 54. Over-fit tree created using all metrics to split days with dry or wet deep slabs or days without deep slabs.

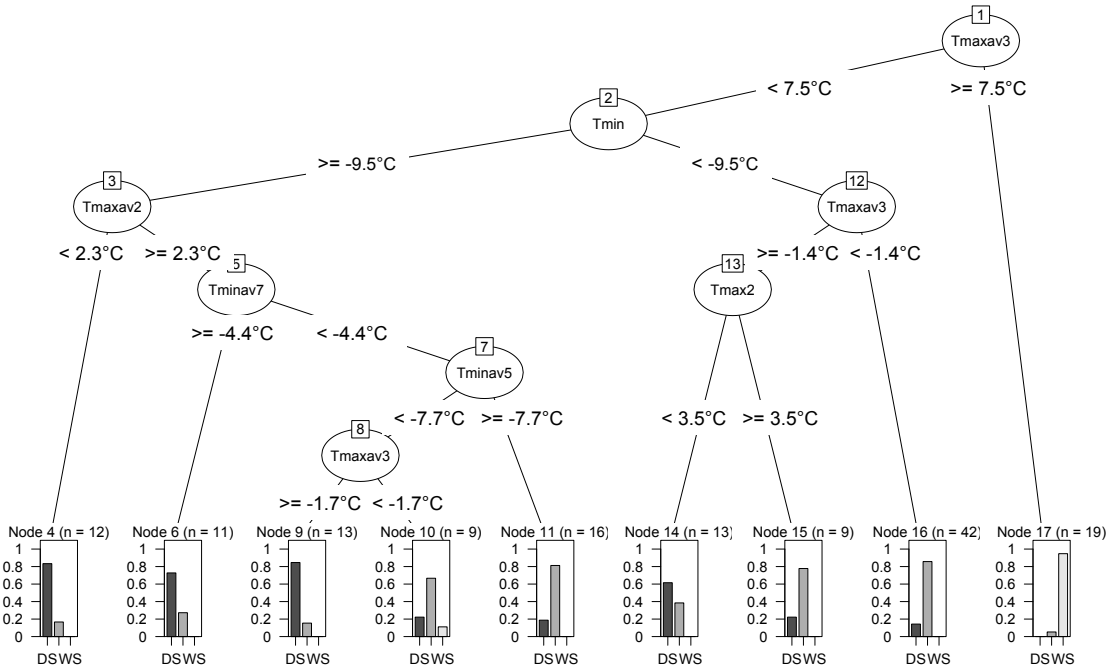



Figure A 55. Over-fit tree grown using only significant metrics to split days by type of deep slab.