TESTING FRACTURE PROPAGATION PROPENSITY FOR SLAB AVALANCHE FORECASTING

CAMERON ROSS **JANUARY 2010**

Applied Snow and Avalanche Research
ASARC
UNIVERSITY OF CALGARY

UNIVERSITY OF CALGARY

Testing Fracture Propagation Propensity for Slab Avalanche Forecasting

by

Cameron Kinnear Houston Ross

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTERS OF SCIENCE

DEPARTMENT OF CIVIL ENGINEERING

CALGARY, ALBERTA

JANUARY, 2010

Abstract

The Propagation Saw Test (PST) is a recently developed field test for slab-avalanche forecasting. It is designed to test the fracture propagation propensity of a buried weak layer and overlying slab. This thesis reports on two winter field seasons of experimental validation of the PST in the Columbia Mountains of British Columbia, in which nearly 800 PSTs, 200 Extended Column Tests (ECT), and 230 experimental short-scaled PSTs were performed. The PST was found to be efficacious at predicting propagation propensity on the slope scale, with predictive skill often exceeding that of other standard snowpack tests. Compared to the ECT, the PST performed well in deeply buried weak layers where fracture initiation via surface loading was difficult. In deep slabs, where field validation was impractical, PST results were compared to forecaster's expert ratings of propagation propensity. The PST was compared to regional avalanche activity and often appeared indicative of propagation propensity trends on the regional scale.

Acknowledgements

I want to extend a foremost thank you to my parents, Don and Evelyn Ross, for supporting my decision to return to school, and for their endless morale and financial support throughout all my academic pursuits. Without them, I'd be miserably packing sewer-line backfill with a jumping-jack compactor. I'm also incredibly grateful to Bruce Jamieson for giving me the chance to study in this incredible field and for all his guidance and feedback during research development and the writing process. Thanks to Dave Gauthier, who moulded this research through his years of pursuing a field propagation test, and whose advice and guidance is reflected in my own work.

For their tireless work in the field, I thank my fellow ASARC students Mike Smith, Thomas Exner, Catherin Brown, and Cora Shea; and technicians Ali Haeri, Mark Kolasinski, Willy Rens, and Jordan Stiefvater in Blue River. Additional thanks to Bruce McMahon, Jeff Goodrich and the entire Avalanche Control Section of Glacier National Park, as well as to Mike Wiegele and the entire staff at Mike Wiegele Heli Skiing, for their exceptional support of the ASARC fieldwork program. Thanks as well to Mike Rubenstein of Kicking Horse Mountain Resort, and to Dale McKnight and Isabelle Thibeault of Chatter Creek Cat Skiing for supporting additional fieldwork.

For financial support of the ASARC program, I would like to acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC), the Helicat Canada Association, the Canadian Avalanche Association, Mike Wiegele Helicopter Skiing, the Canada West Ski Area Association, Parks Canada, the Association of Canadian Mountain Guides, the Backcountry Lodges of British Columbia Association, the Canadian Ski Guide Association, and Teck Coal. I would also like to personally thank Alliance Pipeline for supporting me with the Naomi Heffler Memorial Scholarship in both years of my study, and to thank Bruce Edgerly, Steve Christie and Felix Camire of Backcountry Access for supplying me with top quality field gear.

For on-demand, honest and constructive feedback and proofreading during the writing process, I am grateful to Bruce Jamieson, Dave Gauthier Mike Smith, Cora Shea, Thomas Exner and Katherine Johnston.

I am thankful to my examination committee members Dr. Bruce Jamieson, Dr. Nigel Shrive, Dr. Shawn Marshall and Dr. Lynne Cowe Falls for their thoughtful questioning and thorough review of this thesis.

During the two seasons of winter fieldwork, I had the fortune of learning from many great individuals in the Avalanche Community, including: Bruce McMahon, James Blench, Jeff Goodrich, Johann Schleiss, Brian Gould, Dave Healey, Ron Simenhois, and Brian Robinson. Our discussions helped formulate my respect and understanding of the avalanche phenomena and inspire my research.

Finally, a huge thanks to my little brother, John, for coming out west to play on the mountain-bike trails with me when I needed a break from all this...

Table of Contents

List of Tables

List of Figures

Nomenclature

ACS	Avalanche Control Section (of Glacier National Park)
ALP	alpine (elevation band)
ASARC	Applied Snow and Avalanche Research (University of Calgary)
BTL	below tree line (elevation band)
CAA	Canadian Avalanche Association
CAC	Canadian Avalanche Centre
FPZ	fracture process zone
Fracture Initiation	Weak-layer fractures initiated by an external energy source such as tapping or cutting in a test, or by a skier or explosive etc.
Fracture Propagation	A self-advancing fracture driven by energy supplied by the snowpack and gravity rather than from an external source
g	acceleration due to gravity
HS	total snowpack depth
InfoEx	CAA Industry Information Exchange system
OGRS	Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches (published by the CAA, 2007)
SZ	avalanche start zone
site-layer	an individual layer that fractured during an event at a validation site
TL	tree line (elevation band)
Triggerability	qualitative indicator of the ease of initiating a fracture in a buried weak layer, typically referring to a skier (as in skier-triggering)
WLC	weak layer collapse
Ψ	slope angle
Weak Layers:	
CR	crust

FC	facets
PWL	persistent weak layer
SH	surface hoar
Hand-hardness:	
F	fist (+/-)
4F	four-finger (+/-)
1F	one-finger (+/-)
P	pencil (+/-)
K	knife (+/-)
Snowpack Tests:	
CT	compression test
DTT	deep tap test
ECT	Extended Column Test
FRT	fracture resistance test
PST	Propagation Saw Test
RB	rutschblock test
Score	Index of the effort required to initiate a fracture in a test column via surface loading (equivalent to 'taps' in CT and ECT)
B	break (fracture character)
NR	no result
PC	progressive collapse (fracture character)
RP	resistant planar (fracture character)
SC	sudden collapse (fracture character)
SP	sudden planar (fracture character)
Q1	fast, clean shear (shear quality)
Q2	average shear (shear quality)
Q3	rough, irregular shears (shear quality)
PropL	test result indicating propagation on surrounding slopes is likely

PropUL	test result indicating propagation on surrounding slopes is unlikely

Propagation Saw Test:

x	length of PST column cut at the onset of propagation
y	total length of the isolated PST column
z	vertical depth in the column from snow surface to weak layer
l_p	propagation length
END	propagating fracture reached the end of the test column
SF	propagation stopped in column at slope-normal slab fracture
ARR	propagation stopped at a point of self arrest within test column
yymmdd	weak layer identification (ID) based on date of burial

Extended Column Test:

ECTPV	fracture propagates across the entire column during isolation
ECTP	fracture initiates and propagates across the entire test column
ECTN	fracture initiates but *does not* propagate across the test column
ECTX	no fracture initiates in the test column in the 30 standard taps
n	number of taps (score) when a fracture initiates in the ECT

Analysis and Results:

AAI	Avalanche Activity Index
i	avalanche size class (1-5) used in calculating AAI
N_i	number of avalanches of size class *i*
n	sample size
a	number of correctly predicted events (hits)
b	number of correctly predicted non-events (correct negatives)
c	number of missed events (misses)
d	number of false alarms
POD	probability of detection (sensitivity)
POFD	probability of false detection (specificity)

PON	probability of a null event
TSS	True Skill Statistic
UAA	Unweighted Average Accuracy
E	equally likely (expert rating of propagation propensity)
L	likely (expert rating of propagation propensity)
U	unlikely (expert rating of propagation propensity)

1. INTRODUCTION

This introductory chapter is intended to present the research problem through a discussion of the avalanche phenomena including their impacts on society, their cause and formation, and how practitioners in the field use tools to predict and forecast avalanche conditions. A variety of field tests used to study snowpack conditions associated with avalanches are introduced and discussed, outlining the need for a test specifically targeting the fracture propagation phenomenon that occurs within the failure layer during avalanche release. The Propagation Saw Test (PST) is introduced as a recently developed snowpack test for assessing propagation potential, and the current research objectives are described for refining, validating, and establishing the PST as a practical and effective snowpack test for use in the field.

1.1 Snow Avalanches

Avalanches of rapidly flowing masses of snow descending mountain slopes present annual hazards to people and industry residing in and traveling through mountainous terrain in the winter. These snow avalanches range in size and destructive power from small and relatively harmless to humans, to massive, with the capability of annihilating large swaths of mature forest, destroying buildings, infrastructure, and all human life in their path. It is estimated that snow avalanches cause tens of millions of dollars in losses to industry annually in Canada, including destroyed property and product, as well as delayed just-in-time delivery when avalanche control efforts close transportation routes and debris is cleared from roads and rail lines (e.g. McClung and Schaerer, 2006). As a preventative measure, almost all Canadian transportation routes threatened by avalanche paths have some form of costly monitoring and mitigation program in place (Campbell et al., 2008). In terms of human life, a current average of about 12 people are killed annually in Canada (McClung and Schaerer, 2006), up from the eight person average reported by Jamieson (1995). Campbell and others (2008) estimate that 702 avalanche fatalities have occurred in Canada since the earliest recorded incident in 1782. The past winter of 2008-2009 was particularly deadly with 26 fatalities in Canada and an additional 28 in the United States (avalanche.org).

Historically, large destructive avalanches typically impacted mountainous settlements, mining and logging operations and transportation corridors unexpectedly. As experts became more aware of the avalanche phenomenon and better understood their causes and range in size and destructive potential, the type of avalanche incidents involving humans have changed. Residential development and industrial operations are now designed and built to avoid the runout of avalanche paths, while transportation routes, uninhabited structures and controlled recreation areas are protected through active or passive measures of avalanche control (McClung and Schaerer, 2006 pg. 266). The primary casualty of avalanche incidents in Canada nowadays are unguided backcountry travellers recreating in the vast, uncontrolled mountainous terrain of Western Canada and parts of Quebec and Newfoundland. Beyond the semi-controlled environment of populated communities, workplaces, transportation corridors, ski resorts and mechanized guiding operations, backcountry recreationists expose themselves to varying degrees of avalanche risk, often unknowingly. However, through education and experience, these backcountry users can gain the ability to knowingly and deliberately reduce their exposure risk. Many of the same observations, field methods, test and tools that practitioners in the avalanche field use to reduce or eliminate avalanche risk can be used by backcountry recreationists as well. Following a short discussion of snowpack characteristics typical of North American climates and a brief introduction to the mechanics of avalanche formation and release, many of these field observations and testing methods are described later in this chapter.

1.2 Snowpack

The blanket of snow that is deposited and accumulated over the ground during the winter season is called the snowpack, and, with the exception of glacial ice, is entirely different from one winter to the next. The properties of snow falling from the sky during a storm are the product of numerous meteorological and atmospheric factors including moisture content, temperature, and wind (e.g. McClung and Schaerer, 2006). This, along with metamorphic changes in the deposited snowpack, lend to the development of a layered stratigraphy that exhibits significant spatial variability across regions, drainage basins, and even individual slopes.

Once a new layer of snow is deposited on the ground, metamorphic changes to the crystalline snow can begin instantly and may persist over the duration of the winter creating a layered snowpack that is dynamic throughout the season. These changes largely depend on the original crystal type, the overburden (load) of subsequent snowfalls, and the vertical temperature gradient that exists between the (typically) warm ground and the cold air above the snowpack surface. Metamorphic changes are accelerated when that temperature gradient is high. For example, when the warm ground is buried by an early season snowfall followed by extremely cold air temperatures, a high gradient (e.g. > 10^{o}C/m) develops through the thin snowpack promoting upward water vapour transfer and the kinetic growth of large, angular crystals called facets (Section 1.3.5), which weakens the snowpack (e.g. G. Johnson, 2000). On the other hand, a low gradient with slower vapour transport will favour equilibrium (rounding), which preferentially deposits water vapour in concavities to reduce surface-to-volume ratios, forming rounded equilibrium grains, and generally strengthening the snowpack (McClung and Schaerer, 2006, pp. 53). Future weather events such as rain, wind, or intense solar radiation can also have a large effect on metamorphic process in the near-surface snow (e.g. Bakermans, 2006).

Another factor that affects snowpack metamorphism is the existence of buried melt-freeze crusts which inhibit movement of water vapour through the snowpack. Crusts in the snowpack are known as persistent interfaces (often mistakenly called persistent weak layers) and not only affect snowpack processes but can repeatedly act as a failure surface for avalanches throughout the winter (Section 1.3.4). Other specific failure layers can exist within the seasonal snowpack with various methods of formation and life spans, some of which are described below in Section 1.3.

Although the development of a layered snowpack is common to anywhere where snow falls and accumulates, different regions or mountain ranges can have snow conditions loosely categorized based on their dominant climate (McClung and Schaerer, 2006 pg. 22). On the west coast of Canada and the United States, the snow climate is typically moist and heavy, with frequent and large snowfall events. Mild air temperatures and deep snowpacks result in low temperature gradients and limited kinetic growth (faceting), developing strong snowpacks. Melt-freeze crusts may form and persist through the winter but storm-snow avalanches are the more common avalanche hazard, particularly during and shortly after a storm before the new snow settles and bonds. Conversely, in the drier, colder interior ranges of western North America (e.g. Rocky Mountains), low snowfall and strong temperature gradients promote a thin, faceted and typically weak snowpack. In between the maritime climate of the Coast Ranges and the continental climate of the Rocky Mountains is what is often termed the ‘transitional’ snow climate of the Columbia Mountains in interior British Columbia (BC). This climate is defined by a mixture of moderate temperatures and heavy snowfall events with long periods of cold and clear weather, creating a deep, dry snowpack with a variety of persistent and non-persistent weak layers and interfaces (Hägeli and McClung, 2003; McClung and Schaerer, 2006).

1.3 Weak Layers and Interfaces within a Seasonal Snowpack

The layered snowpack that develops and changes throughout the winter may contain any number of weak layers or interfaces at any time (Figure 1.1). All dry-slab avalanches release on some form of weak layer or interface (McClung and Schaerer, 2006) and they are thus the target of study when digging snow-profiles or performing snowpack stability tests. Weak layers usually have a visible thickness and are of lower strength than the layers above and below, while weak interfaces form the boundary between two layers and may not always be visible in a profile. Weak layers and interfaces are commonly described as either non-persistent or persistent based on their typical life-span in the snowpack. Non-persistent weak layers (e.g. stellars, decomposing fragments) often stabilize within a few days of formation (Jamieson, 1995; Brown and Jamieson, 2006) as metamorphic and pressure-sintering processes form bonds between the adjacent layers. Conversely, persistent weak layers (PWLs) such as surface hoar or facets can remain in the snowpack for weeks or even months and provide a long-term avalanche hazard. A review of 93 fatal slab avalanches in Canada between 1972 and 1991 (Jamieson and Johnston, 1992) estimated that 91% of the fatal avalanches involved failure on a persistent weak layer. A variety of weak layers and interfaces pertinent to this study are described below.

1.3.1 Storm interfaces

When new snow falls during a snowstorm, it often has a consistent grain type, size and moisture content that creates a unique new layer when deposited on the surface of the existing snowpack. This new *storm layer* can be relatively low density and dry compared to the underlying layer, or conversely may be denser with higher moisture content. In either case, these differences in grain type and density create a storm *interface* between the two layers, which can act as a failure plane for avalanches. It is not uncommon for

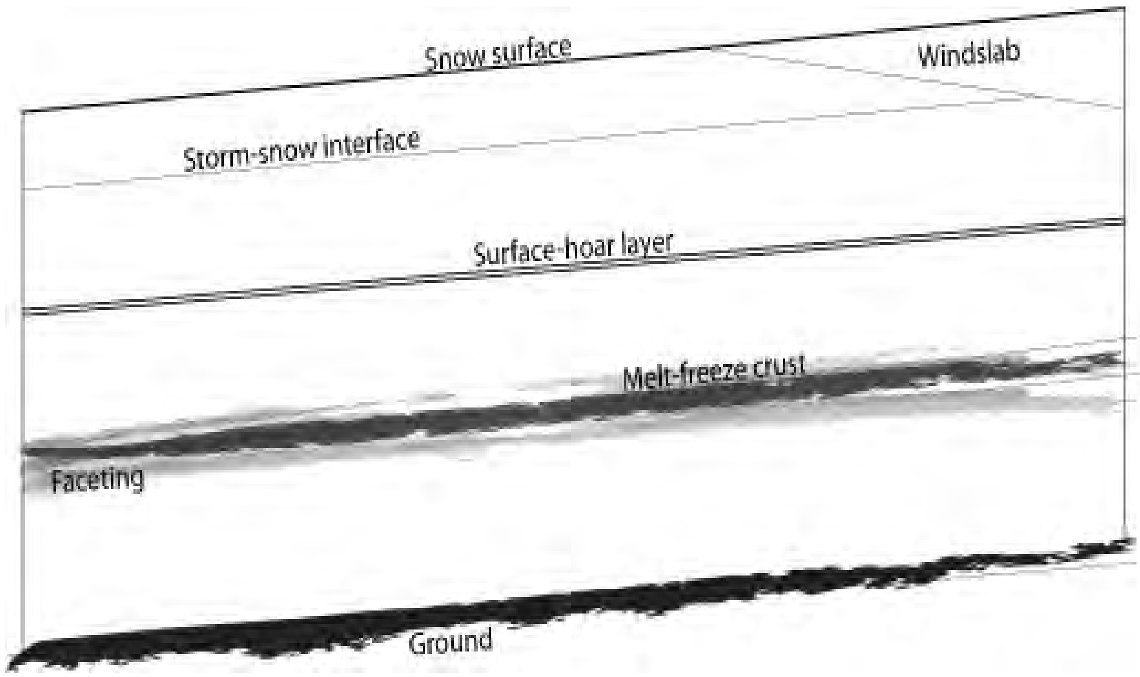

Figure 1.1: Weak layers and interfaces commonly found in the transitional snow climate of the Columbia Mountains of interior British Columbia.

multiple interfaces to develop during the same storm event when atmospheric and weather conditions change mid-storm. Storm interfaces are not always visible in the snowpack and do not typically last very long (non-persistent) due to the processes of settlement, metamorphism, and pressure sintering that occur (often quickly) in the newly deposited snow (Brown and Jamieson, 2006). This is why a large number of natural avalanches occur during and within a few days of a storm before the new snow has time to settle and bond to the sub-layer. Storm snow avalanches mostly slide on the storm interface but may also trigger more deeply buried persistent weak layers (described below) due to the added snow load. Storm snow avalanche sizes are a product of the amount of fallen storm snow and are often of lower density and have minimal cohesion between grains. Thus they are usually classified as *loose* avalanches (Section 1.4), with the exception of wet-snow avalanches or wind-slabs developed during a windy storm.

1.3.2 Wind-slab interfaces

Wind-slabs form during moderate to high wind events when sufficient light, loosely-consolidated snow exists for wind-transport (e.g. McClung and Schaerer, 2006 pg. 32). Wind-slabs are commonly found on the lee side of ridges or small terrain features, where snow scoured from the windward side is deposited as the wind loses velocity over the ridge or terrain feature. This wind-transported snow often deposits in a dense, consolidated slab of small, broken and rounded grains on top of the light, loosely-consolidated snow sheltered from the wind on the lee side, creating a potential failure interface between them. The size and location of wind-slabs are often difficult to identify or predict due to the unpredictable nature of wind.

1.3.3 Surface hoar

Surface hoar, or 'hoarfrost' (Figure 1.2), is the winter equivalent of dew and forms during clear, cool nights when outgoing radiation cools the snow surface, causing the warmer, saturated air above to deposit water-vapour molecules on the surface, growing fragile crystalline feathers (e.g. Lang et al., 1984; McClung and Schaerer, 2006 pg. 49). If the conditions persist over multiple nights of growth with limited destructive solar radiation or wind during the daytime, large crystals may form, particularly on

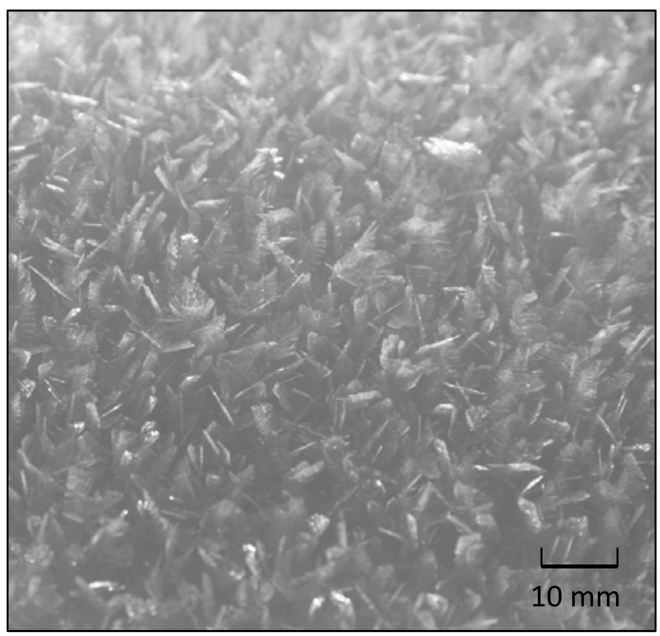

Figure 1.2: Hoarfrost on the surface prior to burial. Once buried, surface hoar becomes a persistent weak layer that commonly fails in destructive slab avalanches.

shady or sheltered slopes (e.g. Breyfogle, 1986). When buried by subsequent snowfall, well developed surface hoar can generate a substantial avalanche hazard and challenge to forecasters. Surface hoar is a primary example of a PWL since its large crystal form is resilient, allowing it to persist in the snowpack for the majority of the season. Surface hoar layers in the snowpack are usually easily visible in a snow-profile and are typically between a few millimetres and a few centimetres thick (Jamieson and Schweizer, 2000).

1.3.4 Crusts

Crusts (Figure 1.3) form on the snow surface and can be the result of rain events or solar or temperature induced melting and refreezing. Often termed ‘rain crusts’, ‘sun crusts’ or simply ‘melt-freeze’ crusts, they aren’t a weak layer themselves, but depending on the length of time and conditions under which they refreeze before a subsequent snowfall, their upper boundary can act as a weak interface and repetitious sliding surface for avalanches (e.g. Jamieson and Johnston, 1992). In addition, facets (Section 1.3.5) are commonly associated with crusts, particularly as the acting weak layer above or below the crust (e.g. Jamieson, 2006c). Because a crust within the snowpack blocks vapour transfer between lower and upper areas of the snowpack (forcing deposition) during strong temperature gradients (e.g. Adams and Brown, 1983; Colbeck, 1991), kinetic growth is promoted, often developing large, weak facets immediately below the crust (Moore, 1982; Fierz, 1998). Also, facets commonly form above crusts, although this process is less well-understood (e.g. Armstrong, 1985; Colbeck and Jamieson, 2001). Most avalanches associated with crusts fail where stress is concentrated in a weak layer (typically facets) above the harder crust (Schweizer and Jamieson, 2001; Schweizer et al., 2003a) and the crust itself acts as the sliding bed surface, not the actual failure layer.

Figure 1.3: An observer pointing at two buried melt-freeze crusts with associated faceting. Common to the Columbia Mountains of interior BC, buried crust-facet combinations create the failure layer in many large avalanches and, along with other persistent weak layers like buried surface hoar, are often the focus of snowpack studies and stability tests.

1.3.5 Facets

Faceting is the result of a strong temperature gradient (e.g. > 10^{o}C/m) within the snowpack that increases water vapour transport (typically upward), driving the kinetic growth of large, square-edged, flat-faced grains (e.g. Colbeck et al., 1990; Miller et al., 2003). Kinetic growth involves the sublimation of water vapour from the top of one crystal which then deposits on the bottom of another (Armstrong, 1985), forming large, angular facets at the expense of smaller, rounded crystals. Facets have minimal cohesion and can be described as resembling 'granular sugar' or as 'rotten' snow.

Facets typically form around crusts (Figure 1.3) where temperature gradients are high and vapour transport can be blocked, concentrating kinetic growth (e.g. Jamieson, 2006c). They also commonly form in the surface snow during long periods of cold temperatures and no snowfall (e.g. Birkeland et al., 1998), or at the bottom of thin snowpacks (basal faceting), where the warm ground creates a strong temperature gradient through the snowpack to the cold air above (e.g. LaChapelle and Armstrong, 1977; McClung and Schaerer, 2006 pg. 57). This latter reason is why faceting is the main snowpack weakness in the colder, drier climate of the Rocky Mountains. If faceting continues for long periods of time, striated facets of multi-layered growth called depth hoar (Colbeck et al., 1990) may form, further weakening the snowpack.

1.4 Avalanche Release

Avalanches are typically categorized as either *loose* or *slab* avalanches, with the latter being either wet-slab or dry-slab. Loose avalanches start as a point release analogous to the rotational failure of cohesionless soil slopes (e.g. Perla, 1980), typically in non-cohesive storm snow on steep slopes or cliffs. As the loose snow slides, it entrains more loose snow and spreads out laterally, generating an avalanche. Loose snow avalanches are generally smaller and lighter and do not have the destructive potential of slab avalanches.

Dry-slab avalanche release begins with increased strain rates leading to fracture in a buried weak layer beneath a cohesive slab of snow (e.g. Schweizer et al., 2003a). The increased strain rate can be the result of additional load from new snowfall, snowpack warming, an explosive shock wave, or the added weight of a person or vehicle. Stress and strain are concentrated at the upper or lower boundary of the weak layer and around flaws in the weak layer, which can be calculated as the ‘stress intensity’ (applied stress energy) in the weak layer and has a critical value known as ‘fracture toughness’ (resistance to fracture). In fracture mechanics, the ratio of fracture toughness over stress intensity determines whether a fracture initiates *and* propagates to release an avalanche (e.g. McClung and Schaerer, 2006 pg. 80). Once started, the initial failure can spread as a self-propagating fracture within the weak layer until a combination of factors cause the overlying slab to fracture in tension at the top (crown) and shear along the flanks, releasing the avalanche (Figure 1.4). The slab can be comprised of multiple snow layers and often consists of a loose storm layer on top of a dense, hard, settled layer which in turn rests on top of the failure layer. Slab avalanches are classified on a scale of 1-5 based on their size and destructive potential, from frequent, small and relatively harmless soft-slabs (size 1) to rare, destructive and deadly deep-slabs (size 5).

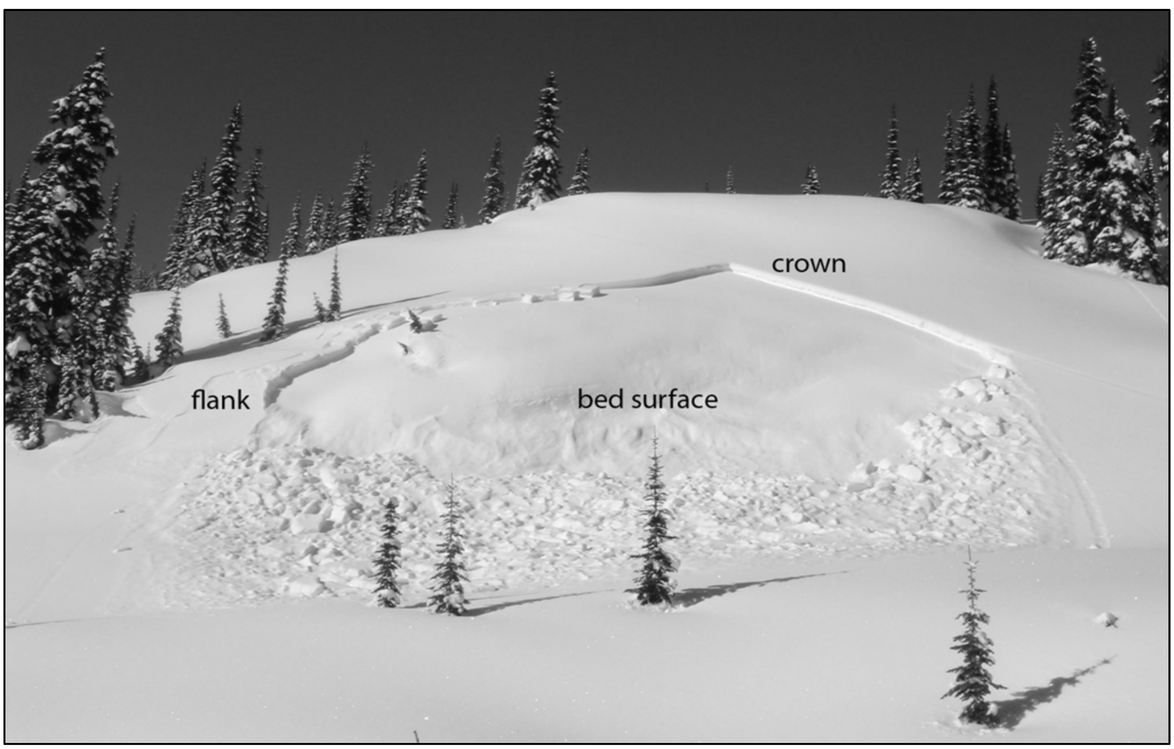

Figure 1.4: A small dry-slab avalanche triggered remotely by a skier a short distance away. The initial weak layer failure under the skier propagated up-slope until a combination of steep slope angle and favourable snowpack characteristics released the avalanche.

In simple terms, the factors that determine the ease of fracture initiation, the extent of propagation and the eventual slab avalanche release include the type and size of terrain feature, the properties of the weak layer, the characteristics of the slab itself, and the hardness and frictional resistance (slope angle) of the bed (sliding) surface. These factors and snowpack characteristics affecting avalanche release are described in further detail in Chapter 2.

In the case of natural slab avalanches, rather than being triggered by the concentrated load of a skier or vehicle, an avalanche occurs under the distributed load of a new snowfall or occasionally under no apparent new load at all. Large numbers of natural avalanches frequently occur during storms in what is termed an 'avalanche cycle'. It is generally agreed upon (e.g. McClung, 1979; Schweizer, 1999; McClung and Schaerer, 2006) that during the formation of these natural dry-slab avalanches, stress concentrations within natural flaws in the weak layer increase strain rates and micro-scale bond breaking causing further weakening of the weak layer over a small area. Eventually, this *deficit zone* can reach a critical crack length and begin to self-propagate. Provided the energy driving propagation exceeds the fracture resistance of the weak layer, the fracture will continue to propagate, often for hundreds of metres if the weak layer is continuous and sufficient driving energy is sustained. The source of this energy is currently the topic of research and debate, and is reviewed briefly in Chapter 2.

Sometimes when a fracture initiates and propagates in a buried weak layer, an audible *whumpf* sound is heard as the layer collapses and propagates outwards but no avalanche is released, especially on flat or low-angle terrain. This rapid downward displacement of the slab is called a 'whumpf' (e.g. Jamieson and McClung, 1996) after the distinctive sound it makes, and is a clear sign of instability in the snowpack even though an avalanche does not release. Sometimes following a whumpf, the fracture can propagate up-slope to steep enough terrain to release an avalanche in a condition called remote triggering (e.g. B.C. Johnson, 2000) such as occured in the avalanche shown in Figure 1.4.

Springtime conditions or mid-winter rain events can introduce melt-water to the pore spaces causing wet-slab avalanches which are beyond the scope of this study. Although wet-slab avalanches are not uncommon, they cause very few fatalities compared to dry-slab avalanches, which, for example, caused an estimated eighty-seven percent of fatal avalanches between 1972 and 1991 (Jamieson and Johnston, 1992). Thus, dry-slab avalanches are the focus of most research, including this study. When the term avalanche is used through the remainder of this study, it is referring to *dry-slab* snow avalanches typical of the mountain ranges in North America during the winter months.

1.5 Growth and Decay of Avalanche Activity on Persistent Weak Layers

As a persistent weak layer such as surface hoar or a crust-facet combination is buried by subsequent snowfalls and a slab begins to develop over time on top of the layer, there is a growth and eventual decay of avalanche activity on that layer, idealized as 'most avalanches' in Figure 1.5. This process can take one to several weeks or months depending on the rate at which the layer is buried and its persistent nature in the snowpack. Generally, the shallow, soft layer of new snow that first buries the layer allows for easy fracture initiation by an external surface load such as a skier, but the new snow rarely possesses the slab characteristics necessary to facilitate extensive propagation of that fracture. This produces some small, low-density avalanches that may also occur naturally under the weight of the new snow. Typically, however, fracture initiation occurs within the weak layer *without* subsequent propagation and no avalanche results. As subsequent snowfalls add to the developing slab, triggering a fracture in the weak layer becomes increasingly difficult under the thicker, stiffer slab. However at the same time, the propensity, or likeliness, of that fracture to propagate *if initiated,* grows since many of the same properties hindering initiation (thick, stiff slabs) favour propagation (e.g. B.C. Johnson, 2000; van Herwijnen and Jamieson, 2007b). This creates a peak in sizeable avalanche activity which is defined by frequent initiation and substantial propagation, followed by a decline in frequency but increase in size (i.e. deep-slab avalanches) as initiation ease continues to decrease and propagation propensity continues to grow.

Eventually, as the weak layer begins to degrade in the snowpack, triggering remains difficult and the propagation propensity also decreases, potentially terminating the active period of that avalanche layer. In some cases, particularly with smooth, early season crusts, the thawing snowpack in spring may again enable easy initiation, effectively re-awakening avalanche activity on that layer if propagation propensity has remained high.

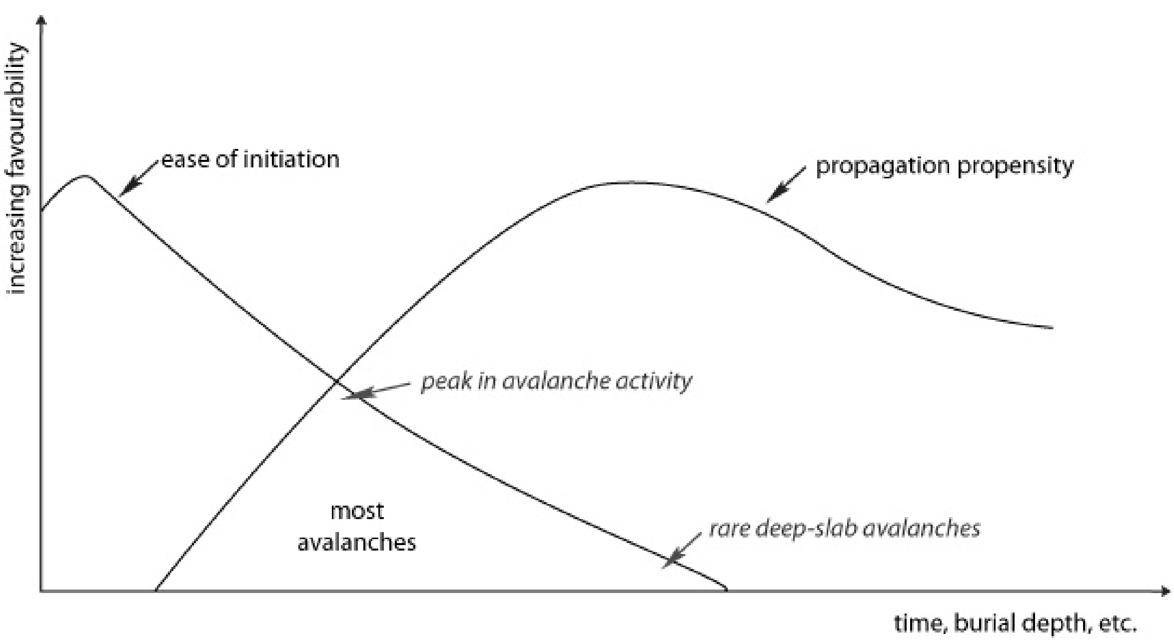

Figure 1.5: A graph roughly depicting the growth and decay of avalanche activity on a specific weak layer. The initial ease of triggering a fracture in a recently buried weak layer becomes more difficult as the layer is buried more deeply by subsequent snowfalls. In some extent of overlapping time, an increase in propagation propensity within that layer as the slab thickens and stiffens creates a peak in avalanche activity where triggering is still common and propagation propensity is high. Eventually, propagation propensity begins to wane as the weak layer degrades in the snowpack (after van Herwijnen and Jamieson, 2007b*).*

1.6 Avalanche Forecasting

The art and science of predicting when and where avalanches may occur and how big they may be is called avalanche forecasting. Operational avalanche forecasting takes place in all mountain operations with an active avalanche control program, such as ski resorts, mechanized guiding operations, many transportation corridors and some industrial operations. Operational forecasting is a day to day activity conducted by a lead forecaster within the controlled area on a known set of avalanche paths. Based on their expert assessment of snowpack and weather conditions, forecasters can recommend action such as explosive control or area closures to reduce the avalanche hazard. Forecasting presents a range of challenges depending on weather and snowpack conditions. For instance, forecasting frequent soft-slab avalanches after a snow storm is generally straight-forward relative to predicting rare but deadly deep-slab avalanches. Guides working for helicopter- or cat-skiing operations or leading groups in to the backcountry also forecast avalanche conditions for the safety of their clients. Guides generally forecast conditions for their entire operational tenure but may also predict conditions for particular ski-runs they may visit that day. Forecasting is different than avalanche planning, which describes the initial efforts of mapping paths and runout extents which help to determine appropriate active control programs or static (passive) defences.

Forecasting avalanche risk for unguided recreation is a much broader task that falls partially on diversely funded public forecasting centers such as the Canadian Avalanche Center (CAC) in Revelstoke, BC. Because of the huge areas and possible variety in weather and snow conditions across the forecast regions, public forecasts for recreation are best used as a general guideline or starting point for users. Often called the *Public Avalanche Bulletin* or *Regional Forecast*, they are based primarily on recent and upcoming weather conditions and persistent weak layers common throughout the region. The task of ‘refining’ the Regional Forecast to fit the local or ‘drainage-scale’ conditions that a recreationist will encounter on a typical day lies on guides leading clients, or on individual recreationists in un-guided parties. Decision support schemes such as the Avaluator (Hägeli et al., 2006) or the Avalanche Terrain Exposure Scale (Statham et al., 2006) can aid in this process by directing users through an assessment of locally observed and regionally forecast conditions to arrive at a relatively safe route of travel.

Although it is largely an art in the sense that good forecasting can be the result of following intuition and recognizing patterns based on experience, science also plays a critical role in avalanche forecasting. Aside from analysing the recent, current and forecast weather, the primary scientific approach to daily forecasting is studying the snowpack, particularly through a snow-profile (Figure 1.6) and snowpack tests (Section 1.8). A snow-profile involves digging a vertical wall and taking observations and measurements of the different layers that comprise the snowpack. The Canadian Avalanche Association (CAA) Observation Guidelines and Recording Standards (OGRS) for Weather, Snowpack and Avalanches (CAA, 2007) describes the accepted way to do a snow-profile, including observing and recording the thickness, grain type and size, hand-hardness, and density of each layer, along with a vertical temperature profile recorded every 10 cm. Hand-hardness is measured on a scale of perpendicular penetration resistance that includes fist (F), four-fingers (4F), one finger (1F), pencil (P) and knife (K) in increasing order, with '+' and '-' adjustments for each.

These measurements help the observer gain valuable information about the location and type of weak layers in the snowpack and the properties of their associated overlying slabs. The temperature profile can help determine if any portion of the snowpack is undergoing kinetic growth (faceting). Schweizer and Jamieson (2007) discussed using a checklist when assessing snow-profiles to identify 'Yellow Flags' that could indicate if the snowpack was conducive to producing avalanches. Although extremely valuable to snowpack assessment, it is important to note that all properties and measurements observed in a profile are subject to spatial variability across the slope and beyond.

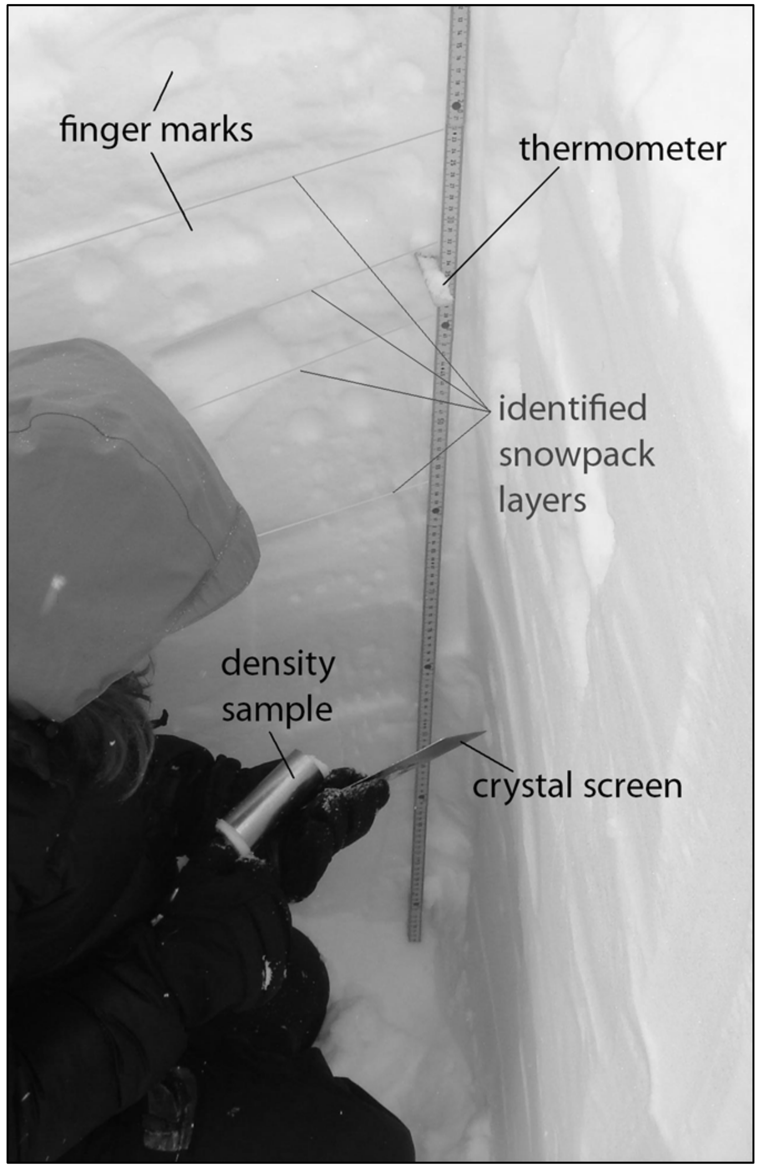

Figure 1.6: A researcher taking a density sample in the midst of a snow-profile. Note the finger marks in the wall from testing the hand-hardness of each layer, and the thermometer next to the ruler used to measure a temperature profile in 10 cm increments.

Additional to the snow-profile, numerous snowpack 'stability' tests exist that enable users to test and observe an isolated sample of the snowpack, typically under some applied surface loading.

Although the profile can provide valuable information about the *structure* of the snowpack and whether that structure is capable of producing avalanches (i.e. slabs and weak layers), dynamic snowpack tests improve the assessment of stability by measuring or 'indexing' the force (energy) required to cause fracture in weak layers beneath a slab, and by allowing observation of the manner or *way* in which that fracture behaves when it initiates. A selection of these tests is described below.

1.7 Testing Conditions for Avalanche Release

1.7.1 Initiation and propagation

As eluded to earlier, the separate but related processes of fracture initiation and fracture propagation are both required for slab avalanche release. Snowpack tests that replicate various proportions of these two processes can reveal the snowpack characteristics and conditions favourable – or unfavourable – to fracture initiation, propagation, or both. This can help researchers develop theories of avalanche release and can aid practitioners to better predict the local avalanche hazard. In artificially triggered (skier, explosive etc.) avalanches, the formation of a crack in the weak layer (fracture initiation) depends on an external energy source generating stress which passes through the slab and concentrates in the weak layer, initiating a propagating fracture if the stress intensity exceeds the fracture toughness. An observer can potentially test this relationship by applying force (external energy) to a sample area of the snowpack to see if fractures initiate and propagate through the test sample, indexing the 'effort' required to do so. In a natural snowpack conducive to avalanche release (not an isolated sample), the fracture will continue to propagate outward from the initial source if sufficient deformation energy is maintained by gravity and the combined characteristics of the slab and weak layer once the initiation energy has dissipated (e.g. Schweizer et al., 2003). Some tests measure and observe the initiation of fractures only, while others, including the test partially developed in this study, observe fracture propagation beyond the initiation stage.

For the purpose of this study, *fracture initiation* is defined as the initial process of crack formation that occurs in the weak layer under dynamic (rapid) loading (artificial trigger) or through slow ductile failure at a natural flaw (natural trigger). In the snowpack tests described below and developed in this study, fracture initiation is induced by the observer (external energy source) through surface loading or by cutting along the weak layer with a saw blade. *Fracture propagation*, on the other hand, is the subsequent process in which the initiated fracture advances rapidly and independently of the initiation energy source and is instead driven by the energy sources inherent in the snowpack itself. For a snowpack test to truly measure or index propagation propensity it must allow for the observation of a self-propagating fracture that is independent of the initiation (energy) source.

A review of the literature describing fracture initiation and propagation in weak snowpack layers, the slab and weak layer characteristics conducive to both, and the snowpack tests that attempt to qualify propagation potential can be found in Chapter 2.

1.7.2 Stable versus unstable slopes

Stable and unstable slopes have been defined slightly differently in various studies (e.g. Jamieson and Johnston, 1993a; Simenhois and Birkeland, 2009; Schweizer and Jamieson, 2003; Gauthier, 2007) and commonly involve a method of skier-testing a slope called 'ski-cutting'. Ski-cutting is the act of skiing across the top of a small (safe) slope and down-weighting at a point along the 'cut' in attempt to trigger a small avalanche or whumpf.

In most studies, 'unstable' slopes are defined as those that avalanched or whumpfed naturally, accidentally, or during explosive- or skier-testing. On the other hand, 'stable' slopes are defined as those slopes that did *not* avalanche or whumpf after repeated skier or explosive testing. Some studies occasionally consider un-avalanched slopes as 'stable', even if not skier or explosive tested (Simenhois and Birkeland, 2009); some include 'shooting cracks' along the snow surface as a sign of instability (Winkler and Schweizer, 2009); and some occasionally rate slopes as stable or unstable based on the results of snowpack tests on that slope rather than skier or explosive testing (Simenhois and Birkeland, 2009; Winkler and Schweizer, 2009). For the purpose of this study, which focuses on the propagation propensity of slab and weak layer combinations on the slope, stable and unstable slopes are defined similar to Gauthier (2007, pp. 89) and Gauthier and Jamieson (2008a): *Unstable* slopes are those that show obvious signs of high propagation propensity observed in a natural, accidental, or skier-cut avalanche or whumpf up to 24 hours prior to testing. *Stable* slopes, on the other hand, are only those slopes that were skier-tested and had the ski-tracks dug up to confirm that a weak layer fractured during the ski-cut, but had not propagated. In this sense, stable slopes are those that show little or no fracture propagation propensity when initiated, and the stable rating *only* applies to the weak layer(s) that fractured during the ski-cut (see Section 3.5). If a fracture did not initiate in a weak layer during a ski-cut, nothing can be said about the propagation propensity of that layer and its overlying slab, and therefore no rating can be assigned (stability unknown).

1.8 Standard Snowpack Stability Tests

Numerous snowpack tests have been developed over the years to aid avalanche practitioners and recreationists in assessing snowpack strength and stability. Most involve some form of applied dynamic surface loading, while others measure weak layer shear strength by means of slope parallel pulling (e.g. shovel shear test, shear-frame test; CAA, 2007). Typically, the ‘effort’ required to cause failure in the test is indexed in the form of a *test score* that can be used to help rate local stability. By establishing a ‘threshold value’ at which the test score interpretation moves from indicating unstable conditions to stable conditions, the local slopes can be rated unstable or stable accordingly. In their development, stability tests are often validated by comparing test results next to skier-tested slopes that have been determined as stable or unstable by various methods (e.g. Jamieson, 1999; Jamieson and Johnston, 1993a; Gauthier and Jamieson, 2008a, 2008b; Simenhois and Birkeland, 2009). If the test result matches the actual stability for the tested slope, it is a ‘correct prediction’. If the results indicates stability when in fact the slope failed, it is a ‘false-stable’ result, while the opposite case is termed ‘false-unstable’. Two standard stability tests pertinent to this research – the compression test and the rutschblock test – are introduced here and methods and recording procedure for both are described in detail in the CAA OGRS (CAA, 2007). Chapter 2 reviews efforts to validate the rutschblock and compression test in terms of evaluating propagation propensity.

1.8.1 Compression test

The compression test (CT) is the preeminent snowpack stability test in Canada, and the first most recreationists learn (Figure 1.7). Developed by park wardens in Banff National Park (Jamieson, 1999), it involves a 30 cm by 30 cm column isolated to about 100 cm deep in the snowpack and is loaded from the surface by placing a shovel blade atop the column and 'tapping' 10 times from the wrist, 10 times from the elbow, and 10 times from the shoulder (CAA, 2007).

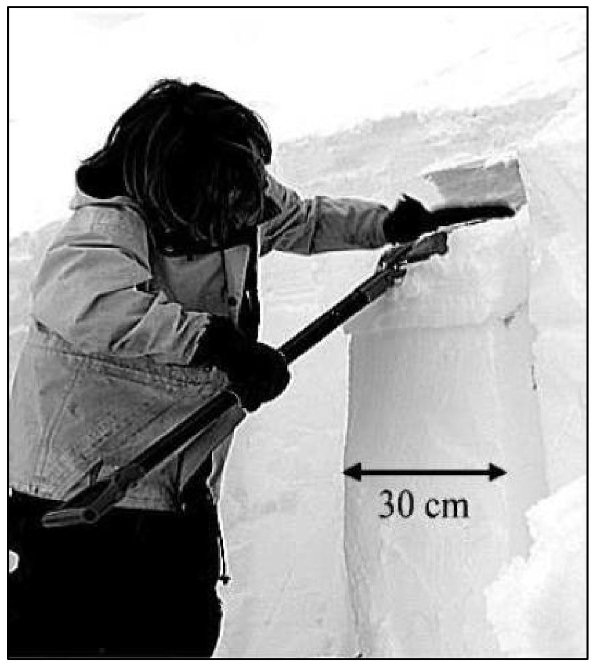

Figure 1.7: The compression test (CT) (Applied Snow and Avalanche Research, University of Calgary (ASARC) photo).

The current number of taps (CT score) when a fracture initiates in the column is recorded as a measure of initiation ease, the depth to fracture is measured, and an observation of the 'fracture character' is taken (e.g. planer, sudden collapse, broken). If no fracture occurs in the 30 taps, 'NF' is recorded and a score of '35' may be given (Jamieson, 1999). Fracture character as it relates to propagation has recently been the focus of research on the CT (e.g. van Herwijnen and Jamieson, 2007a) and is discussed in Chapter 2. Variations of the CT include the rammrutsch (Schweizer et al., 1995b), the stuffblock (Birkeland et al., 1996), and the drop hammer (Stewart, 2002). A variation on the CT for deeply buried weak layers, called the deep tap test (DTT), follows the same method as the CT with the exception that the top of the column is removed to within 15 cm of the weak layer of interest at the back of the column (Jamieson, 2003; CAA, 2007). This enables transmission of surface loading to a weak layer that would otherwise not have 'felt' it from the snow surface.

1.8.2 Rutschblock test

The snowpack test most closely replicating the dynamic impact of a skier on a slope is the rutschblock (RB) test (Föhn, 1987). The rutschblock (Figure 1.8) involves a large (1.5 x 2.0 m) block isolated from the surrounding snowpack and loaded in six progressive steps by a skier (stepping on, knee bends, jumping etc.).

Figure 1.8: The rutschblock test (RB) in progress, on the fifth loading step. The upper weak layer released on an earlier loading step but the lower layer has yet to fail.

When a weak layer failure within the block is observed, the loading step (RB score) is recorded along with a rough measure of the amount of block that released with failure (e.g. whole block, part block). The lower the score and the greater the proportion of block that slid, the greater the instability (Schweizer, 2002). Developed by the Swiss Military for testing weak snowpack layers, the rutschblock has frequently been validated as an effective snowpack test in which low scores correlate well with skier triggered avalanches (e.g. Fohn, 1987; Jamieson, 1995; Jamieson and Johnston, 1993a). More recent research comparing the RB score and the release type independently to observed propagation in the field is reviewed in Chapter 2.

1.9 Propagation-specific Snowpack Tests

Despite the recent efforts to compare RB and CT results to observations of propagation in the field, these two tests depend on the dynamic waves from surface loading being transferred to the buried weak layers, which is known to diminish and disappear with increasing depth in the snowpack (e.g. Schweizer et al., 1995b; Schweizer and Campanovo, 2001). Thus, they are perhaps better measures of the *ease of initiation*, rather than the *propensity for propagation* within that layer (e.g. Gauthier and Jamieson, 2007a). Another limitation to snowpack tests that depend on surface loading is the importance it puts on site selection when performing the tests. Because of the high variability in snowpack depth even on the local scale, selecting a test site that *appears* representative of a ski slope may in fact be greatly different if the unstable weak layer at the test site is much more deeply buried than on the potential ski slopes. In this case, the deep weak layer may not ‘feel’ the transmitted surface load and no fracture will be observed at the test site but, despite the non-initiation result, propagation propensity in that layer may be high and could release a potentially fatal avalanche if triggered from a shallow snowpack spot on the ski slope. Conversely, shallow storm layers often fracture easily under surface loading but do not possess the slab and/or weak layer characteristics conducive to propagation. The traditional focus of snowpack tests on fracture initiation highlights a need to be able to test for weak-layer fracture propagation propensity specifically, perhaps independently from the ease of initiation. Two recent snowpack tests have been in development that attempt to solve this knowledge gap in the snowpack/avalanche forecasting toolbox.

1.9.1 Extended Column Test

The Extended Column Test (ECT) is one such propagation test and is essentially a widened CT that still depends on surface loading to obtain results. Developed in Colorado and New Zealand by Karl Birkeland and Ron Simenhois (e.g. Simenhois and Birkeland, 2006; 2009), the ECT uses an elongated CT column (30 cm up-slope by 90 cm across-slope) and is loaded at one end of the column in the same steps as the CT (Figure 1.9).

The operator takes note of the number of taps required to initiate a weak-layer fracture (ECT score), and the subsequent number of taps required to propagate that fracture across the remainder of the column. The test indicates propagation is *likely* only when propagation occurs on the same or one additional tap as initiation, and only within the 30 standard loading steps (Simenhois and Birkeland, 2007).

Figure 1.9: An operator performing the Extended Column Test (ECT). The ECT was developed to test both fracture initiation ease and propagation propensity by extending the cross-slope length of the CT.

1.9.2 Propagation Saw Test

The Propagation Saw Test (PST), under development at the University of Calgary since 2005 (e.g. Gauthier and Jamieson, 2006b), is the second test and is the primary focus of this thesis. The PST utilizes an isolated column that is 30 cm across slope by 100 cm or more up-slope. Rather than using surface loading to initiate a fracture, the dull edge of a thin saw blade is drawn up-slope through the weak layer of interest (often referred to with the misnomer 'saw-cutting') until the fracture propagates suddenly from the end of the saw-cut towards the end of the column (Figure 1.10). The column length is dependent on the weak layer depth, and is 100 cm long for all layers shallower than 100 cm but is extended to equal the weak layer depth when the layer depth exceeds 100 cm. Gauthier (2007), and Gauthier and Jamieson (2007b) determined that propagation on similar surrounding slopes is only likely when the fracture self-propagates to the end of the column before 50% of the column length has been cut. When the saw cut exceeds 50%, or the propagating fracture stops at a slope-normal fracture through the overlying slab or at a point of self arrest in

Figure 1.10: An operator performing the Propagation Saw Test (PST). The PST was developed to test fracture propagation propensity independently of fracture initiation via surface loading.

the layer, propagation on similar surrounding slopes is said to be *unlikely*. Unlike the other snowpack tests described, the PST makes no inference on the likelihood of *triggering* a weak layer while skiing, only the likelihood of that layer propagating a fracture *if* triggered.

The development and initial validation of both the PST and ECT are reviewed in detail in Chapter 2, and the test methods and recording standards are described with greater detail in Chapter 3.

1.10 Thesis Overview and Objectives

The recent focus on fracture propagation both as an integral part of the avalanche release process (e.g. Heierli and Zaiser, 2006, 2007; McClung, 2009a), and as a subject for field-testing (e.g. van Herwijnen and Jamieson, 2005; Gauthier, 2007) provides the motivation and funding for this research. Both the Propagation Saw Test and Extended Column Test are fairly recent creations with many unanswered questions about their test methods, limitations, and predictive validity. This study attempts to answer some of those questions, primarily regarding the PST. Through comparative analysis with the ECT, expert ratings of propagation, and regional avalanche occurrence data, and through further validation of the PSTs geometry and prediction results, this study aims to assess the efficiency and accuracy of the PST as a reliable predictor of propagation propensity in buried weak layers.

The objectives of this study are as follows:

- To further validate the PST's predictive accuracy next to sites of observed fracture propagation, or confirmed initiation *without* propagation, in the field.

- To experiment with scaling PST column length below 1 m with weak layer depth to determine if the standard method presented by Gauthier (2007) is a more effective and accurate predictor of propagation propensity.
- To compare the test methods and results of the PST and ECT side by side, determining strengths and limitations of each test and the slab characteristics that influence test results.
- To determine if expert ratings of propagation propensity can be used to validate deep-slab PST results where the above validation method is impractical.
- To compare PST results to regional avalanche occurrence data in attempt to validate the PST on the regional scale.

The basic snowpack and avalanche information introduced in Chapter 1 provides some prerequisite background knowledge and introduces the motivation for, and the objectives of this study. A review of the literature is presented in Chapter 2, detailing the development of both the PST and ECT, and describing recent efforts to assess propagation propensity using the traditional snowpack stability tests described earlier. In Chapter 3 the field methods used to collect data for this study are described, followed in Chapter 4 by the results and analysis apropos to the objectives. Chapter 5 summarizes and concludes the study, presents a brief discussion of the PST's application to slab-avalanche forecasting, and suggests avenues of further related research.

2. LITERATURE REVIEW

Over the past decade, researchers have begun to rigorously re-evaluate the physics of propagating fractures in the avalanche formation process – a phenomenon that practitioners and avid backcountry recreationists have been familiar with for decades, even if they have been unsure of the underlying physics. Although a full review of the physical and mechanical theories relating to propagating fractures in snow is beyond the scope of this study, some key ideas and developments leading to the current state of knowledge are presented here (Section 2.1). Of more recent development, and with greater pertinence to this study, is the way we *test* the snowpack for fracture propagation propensity in the field. A critical review of the literature related to field-testing for fracture propagation is presented, covering attempts to gain information on propagation propensity from the existing standard stability tests (Section 2.2), as well the development of the PST (Section 2.3) and ECT (Section 2.4) as propagation-specific snowpack tests. A review of the literature describing snowpack characteristics conducive to propagation (Section 2.5) follows. Short reviews of forecasting studies associated with deep slabs (Section 2.6) and the regional scale (Section 2.7) are included, describing any use of snowpack tests in their methods.

2.1 Key Developments in the Theory of Fracture Propagation in Snow

Beginning ten years ago, field experiments that observed the propagation of weak-layer fractures in snow (B.C. Johnson, 2000; Johnson et al., 2004; van Herwijnen, 2005; van Herwijnen and Jamieson, 2005; van Herwijnen et al., 2008) began to challenge the long-accepted theories of propagating shear fractures (e.g. McClung, 1981, 1987, 2005a; Bader and Salm, 1990), and instead promoted the role of weak-layer-collapse induced bending waves as the driving force behind propagation. These influential field experiments are reviewed in Section 2.1.1 followed by a brief review of the shear fracture mechanics models (Section 2.1.2) and the more recent weak layer collapse models (Section 2.1.3).

2.1.1 Fracture propagation observed in field experiments

Few studies have observed the propagation of weak layer factures in the field, but have proven influential in the development of associated theories and propagation propensity tests. Johnson and others (2004), for instance, used geophones to measure propagation speeds in weak layers on horizontal terrain, and concluded that compressive failure (collapse) of the weak layer on low-angle terrain provided sufficient energy for fracture propagation, and that the velocity of the resulting bending wave in the slab was dependent on the slab stiffness. This complimented an earlier model proposed by B.C. Johnson (2000) that suggested propagating linear flexural waves were the failure mechanism for whumpfs and remotely triggered avalanches from near-horizontal terrain, differing from the shear-propagation mechanism commonly accepted for avalanches on inclined slopes. Although his model was later amended to include a gravitational term (Heierli, 2005), he is credited with contributing to the turning point in propagation research (Heierli et al., 2008b).

More recently, van Herwijnen and colleagues (van Herwijnen, 2005; van Herwijnen and Jamieson, 2005; van Herwijnen et al., 2008) used a high-speed video camera to capture fracture propagation in a variety of snowpack field tests. The earlier studies used particle tracking in a cantilevered beam test similar to B.C. Johnson (2000), as well as the rutschblock and compression test, to capture the fracture process and concluded that slope-normal displacement due to crushing of the weak layer accompanied the propagating fracture, and was independent of slope angle (van Herwijnen and Jamieson, 2005). In their later experiment, van Herwijnen and others (2008) used a 3-4 metre long version of the PST with a similar method to monitor the fracture propagation process and showed that a bending wave propagated through the column associated with weak layer collapse with distinct slope-normal displacement preceding slope-parallel sliding. In separate experiments using an independently developed test similar to the PST, Sigrist and Schweizer (2007) combined field tests with finite-element simulations to determine the critical energy release rate required to propagate fractures, and found that slope-normal bending of the slab contributed considerable energy to the process in addition to slope-parallel shear deformation.

2.1.2 Shear failure models of fracture propagation

For almost thirty years prior to these influential field experiments, the accepted paradigm for natural and artificially triggered avalanche release were models of shear-fracture propagation (Perla and LaChapelle, 1970; McClung, 1981, 1987; Gubler and Bader, 1989; Bader and Salm, 1990; Louchet 2001a, 2001b; Bažant et al., 2003). For natural avalanche release, McClung (1979, 1981, 1987) modeled a shear band initiated at stress concentrations around a natural flaw in the weak layer that slowly expanded through strain softening (ductile failure) at the edges until a critical length was reached and the band rapidly self-propagated as a brittle fracture (Figure 2.1). This process was driven by the slope-parallel component of gravity. McClung's earlier (1977, 1979) laboratory work on snow subjected to slow shear deformation supported the hypothesis of ductile shear failure in the weak layer. Furthermore, Jaccard's (1966) study of stress concentrations at the edges of local weak-layer fractures was the first in a series of studies (e.g. Brown et al., 1973; Smith and Curtis, 1975; Lang and Brown, 1975) surrounding Perla and LaChapelle's (1970) work that supported the concept of local flaws in the weak layer – now commonly called *deficit zones* – as the initial source of failure.

The shear fracture model of McClung (1981, 1987) was essentially an energy balance equation of shear band driving force (stress intensity) versus resistance force (fracture toughness) that promoted weak layer ductile shear failure as the source of primary fracture, and strain-energy near the crack tip (fracture process zone, FPZ) as the energy source for propagation in slab avalanche release. McClung (2002) made early theoretical estimates of the Mode II (in-plane shear) fracture toughness at the base of dry snow slabs and concluded that the fracture mechanics size effect law described for snow by Bažant and others (2003) governed fracture toughness through a scaling law based on slab thickness. Field measurements from slab avalanches were used to support the dependence of fracture toughness on slab depth and weak layer nominal shear strength (McClung, 2005a), and to estimate values of fracture energy (McClung, 2007a) and fracture toughness (McClung, 2005a; McClung and Schweizer, 2006), both within the weak layer (Mode II in-plane shear) and through the slab (Mode I tension). McClung (2005a) argued that fracture toughness, although fundamental to predicting propagation in avalanches, could never practically and safely be measured prior to avalanche release.

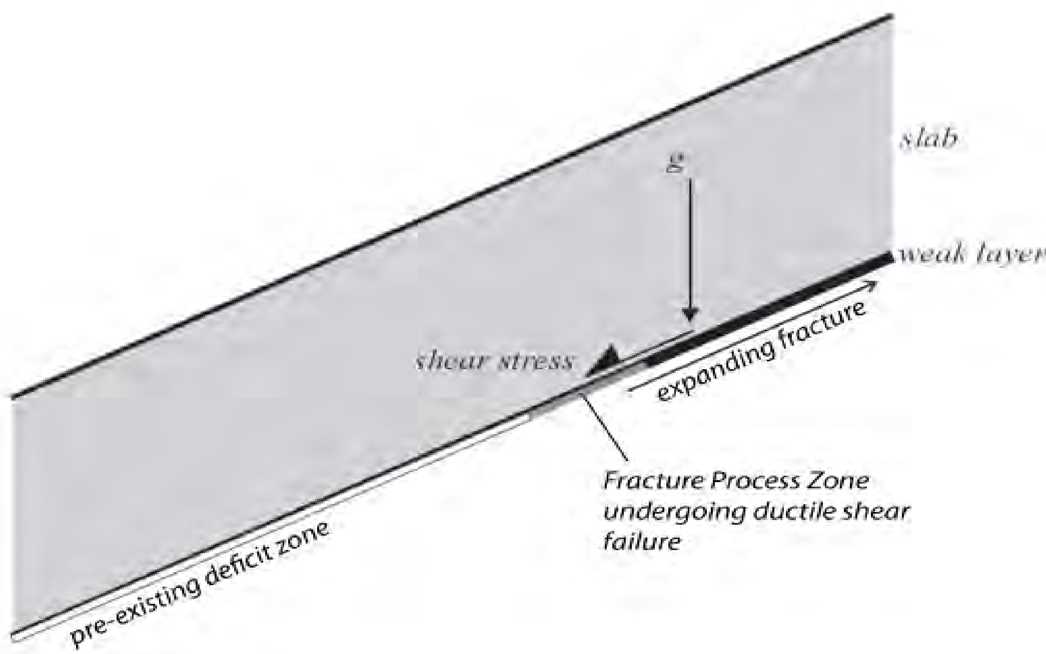

Figure 2.1: Basic schematic of the shear fracture mechanics model developed by McClung (1981, 1987). It models a slowly expanding ductile shear failure at the edges of an existing deficit zone in the weak layer that leads to rapid fracture propagation when a critical length is reached.

Instead, McClung and Schweizer (2006) suggested that the ratio of tensile toughness in the slab (measureable) to fracture toughness in the slab and weak layer system (un-measureable) may be useful in estimating potential avalanche size based on a limited range of slab thicknesses (0.2 – 1 m). McClung (2008, 2009b) showed that slab thickness alone can be used to roughly estimate avalanche size based on field measurements at avalanche sites.

In a review of the shear-propagation models, Schweizer (1999) noted that all were two-dimensional, required an inclined slope to translate the gravitational weight of the slab into slope-parallel shear force, and assumed a pre-existing weakness or flaw in the otherwise homogeneous weak layer that was unable to support the applied shear stress, thereby initiating the ductile failure process. The models succeeded at offering an explanation for natural and artificially released avalanches on slopes, but failed to explain propagation on the flats (remote triggers and whumpfs), and failed to identify the source of the required pre-existing flaws or deficit zones. Schweizer (1999) pointed out the need for a link between micro-scale snow failure and macro-scale slope (slab) failure and implied that such micro-scale damage leading to fracture nucleation (e.g. Nye 1975) could replace the assumption of pre-existing deficit zones used in most models.

2.1.3 Weak layer collapse (WLC) models of fracture propagation

Bucher (1948) and Bradley and Bowles (1967) originally proposed that thick weak layers of depth hoar could collapse, implying the critical shear stress for such layers need not be parallel to the slope as in failure models developed for thin weak layers and interfaces (e.g. McClung, 1981, 1987; Bader and Salm, 1990). Based on field data, Jamieson (1995 pp. 176) argued that slope-parallel shear stress need not be critical for the failure of relatively thin weak layers, including surface hoar. Earlier, Lackinger (1989) had proposed that slab bending associated with collapsing weak layers could propagate fractures.

Schweizer and others (2003a) reviewed the contemporary state of knowledge on avalanche release and made the distinction between the fracture initiation process and the subsequent process of fracture propagation. The review reiterated that the primary failure is between the slab and substratum (either in a weak layer or along an interface), and that it is usually in slope-parallel shear, but occasionally compressive failure could lead to loss of shear support in thicker weak layers. They also concluded that shear is essential for propagation at the scale of slab thickness, but that at the scale of grains and bonds, fracture can be any mode, including mixed-mode (tension, compression, and/or shear).

Complementing the field experiments measuring propagating fractures (Johnson et al., 2004; van Herwijnen and Jamieson, 2005; Gauthier and Jamieson, 2006a), Heierli (2005) and Heierli and Zaiser (2006) introduced a new analytical model based on weak layer collapse (WLC) (Figure 2.2) that could physically and mathematically explain the recent field observations by calculating the required energy release for propagation due to a localized collapse in the weak layer. The WLC model arrived at a similar velocity of approximately 20 m/s reported by Johnson and others (2004) and van Herwijnen and Jamieson (2005), but showed that the velocity could be dependent on slab thickness. The model was later updated (Heierli and Zaiser, 2007) to include energy from slope-parallel shear deformation and showed that failure could be driven by energy contributions from either source, primarily dependent on slope angle. This model effectively closed the gap between slope-triggered avalanches and the propagation across near-horizontal terrain observed in whumpfs and remotely triggered avalanches. Recently, the model was expanded again to explain natural, spontaneous avalanches under no apparent loading through the development of a mixed-mode anticrack theory for fracture nucleation and propagation in a collapsible weak layer (Heierli et al, 2008a, 2008b; Heierli, 2009).

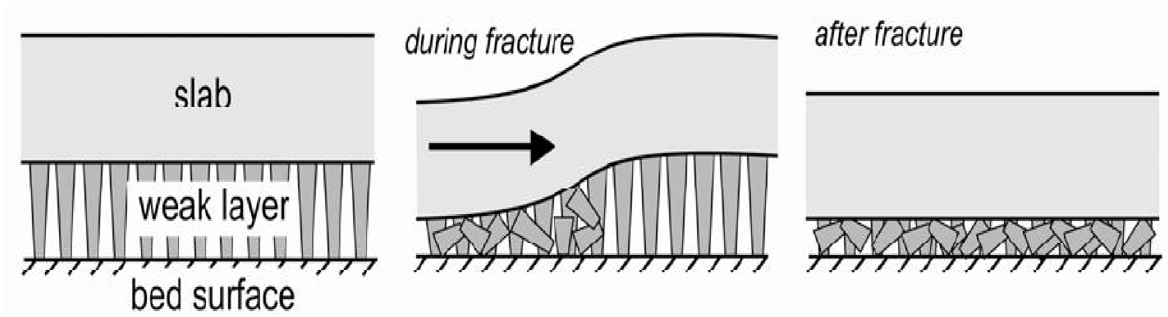

Figure 2.2: The Weak Layer Collapse (WLC) model, showing a collapsing bending wave propagating a fracture along the weak layer (from Gauthier, 2007, with permission).

McClung (2005b, 2007b, 2009a) addressed the observations of collapse in thick (10 mm +) weak layer propagation as a process that occurs at and behind the crack tip (FPZ) of a propagating shear fracture and that energy released by vertical collapse would lead to propagating shear wave rather than a flexural bending wave. For reasonable slab and weak layer properties, propagating shear-fracture velocities could be calculated that matched those reported by Johnson and others (2004) and van Herwijnen and Jamieson (2005). In his most recent work, McClung (2009a) further reconciled the importance of weak layer collapse and noted that, in some cases, shear propagation from slope-parallel deformation alone would be an inadequate explanation. Instead, he proposed that released gravitational energy contributes to the advance of a shear fracture in the FPZ, but the process could still be modeled as an expanding shear crack similar to early models (McClung 1981, 1987).

2.1.4 Summary: towards field-testing for fracture propagation

In their review of snow avalanche formation, Schweizer and others (2003a) posed the following question: "*Which mechanical properties of which slab/weak layers describe the* *propensity for fracture propagation and how can practitioners test for propagation propensity?*"

The recent published observations of propagation in the field (e.g. Johnson et al. 2004; van Herwijnen and Jamieson, 2005) led researchers to not only re-evaluate the current paradigm on weak-layer fracture propagation in avalanche release, but also to explore ways in which they could *test* the snowpack for propagation propensity in weak layer and slab combinations. Although the above discussion is a brief summary of the development of our current knowledge surrounding fracture propagation in weak snowpack layers, the real purpose is to provide a background for the research specifically related to this study: the development of fracture propagation field tests that could aid avalanche practitioners, and perhaps even recreationists, at predicting the propagation propensity of buried weak layers on surrounding slopes. The remainder of this chapter is a critical review of the literature pertaining to testing the snowpack for propagation propensity with intentions of answering the question posed by Schweizer and others (2003a).

2.2 Observations of Propagation in Traditional Snowpack Tests

Early efforts to test the snowpack for propagation propensity had researchers looking to the existing selection of standard snowpack stability tests for answers, primarily the small-column compression test and the large-column rutschblock test. For example, Schweizer and others (2003a) noted that based on the estimated shear deficit size of 0.1 – 1 m required for self-propagating brittle fractures due to rapid loading (Schweizer, 1999), the large-column stability test methods (i.e. rutschblock) did essentially test a comparable sized area. In attempts to gain more valuable information from the compression test, observations of fracture character (e.g. van Herwijnen and Jamieson, 2002) or equivalently shear quality (e.g. Johnson and Birkeland, 2002) were being made in attempt to qualify the propagation propensity of the fractured weak layer. In Switzerland, Schweizer and Wiesinger (2001) developed a system for rating the type of release and the fracture quality observed in the Rutschblock based on an earlier proposal by Schweizer and others (1995a). For almost three decades previous, the CAA's Observation Guidelines and Recording Standards for Weather, Snowpack, and Avalanches (OGRS) had encouraged noting obvious collapse in the weak layer and settlement of the slab when performing a shovel test, while more recent versions of OGRS suggested – without verification – that sudden, planar failures were more indicative of avalanches than less distinct fractures (van Herwijnen and Jamieson, 2002). These references to observing fracture quality in early practitioner guidelines indicate that avalanche workers were aware that the *way* in which the weak layer failed in tests, not just the force under which it failed, revealed something important about the snowpack's behaviour.

2.2.1 Fracture character and shear quality in stability tests

In a review of the compression test, Jamieson (1999) proposed a recording standard for different types of weak-layer fracture character commonly observed in the compression test and other surface-loading snowpack tests, but provided no data to substantiate them. van Herwijnen and Jamieson (2003) refined the original fracture character types proposed by Jamieson (1999) and included the sudden collapse (SC) of the weak layer; a progressive collapse (PC) of the weak layer observed over multiple taps; broken failure (B) in which the fracture was not planar along the layer; and the newly separated sudden planar (SP) and resistant planar (RP) fractures that differentiated between thin planar fractures that were fast and 'popped' and those that were slow and resistant to sliding (Table 2.1).

The same study (van Herwijnen and Jamieson, 2003) related weak layer and snowpack characteristics to the observed fracture character recorded in a year of compression and rutschblock tests under the new classification scheme, and determined if differences in fracture character were associated with avalanche activity. Results revealed that 97% of fractures in persistent weak layers (surface hoar and facets) were either SP or SC. Consequently, and as anticipated, SP and SC results showed a strong association with skier triggered avalanches, whereas RP and PC results were more common on un-triggered slopes. This showed improvement over a previous study based on five years of data under the original classification scheme in which SP and RP results were grouped together (van Herwijnen and Jamieson, 2002), and suggested that fracture character, in addition to test score, could be used to predict skier-triggered avalanches.

More recently, van Herwijnen and Jamieson (2004a, 2004b, 2007a) and van Herwijnen (2005) combined their previous datasets (2002, 2003) under the new classification scheme for fracture character to strengthen their analysis. They arrived at similar results and offered an explanation for the observed link between sudden fracture characters, persistent weak layers and avalanche activity. They observed that SC and SP results were most often associated with deeper, larger grain, persistent weak layers with a stiff layer immediately above that were sometimes difficult to initiate but showed a high propensity to propagate and release avalanches (e.g. Schweizer and Jamieson, 2001). On the other hand, most of the PC and RP fractures were associated with shallow, soft, non-persistent weak layers under less stiff slabs that favoured fracture initiation but were not conducive to propagation (e.g. Schweizer et al., 2003a) and rarely released slab avalanches.

Concurrently in the United States, Johnson and Birkeland (1998) and Birkeland and Johnson (1999) proposed a three level system of classifying ‘shear quality’ during their development of the stuffblock test (Birkeland et al., 1996; Birkeland and Johnson, 1999). They felt the system provided additional important information about slope stability and interlayer relationships. The classifications were Q1 for fast, clean shears; Q2 for what they called ‘average’ shears that were mostly smooth but resistant to sliding; and Q3 for rough, irregular shears (Table 2.1). In 2002, they expanded on their classification system by reporting on six seasons of compression, stuffblock, and rutschblock data next to signs of instability (or lack thereof) in an effort to determine if interpreting shear quality in stability tests would improve snowpack evaluation (Johnson and Birkeland, 2002). Results showed that the Q1 shear quality was strongly associated with signs of instability and they emphasized the condition of when a 'high' score (suggesting good stability) in a snowpack test had a Q1 shear quality as an example of when shear quality could improve the interpretation of the test.

Table 2.1: Fracture Character scheme proposed by van Herwijnen and Jamieson (2003) for use with the compression test, matched to classifications of Shear Quality recorded by Johnson and Birkeland (1998) in the stuffblock test (after van Herwijnen and Jamieson, 2004b). A 'sudden' fracture character or the equivalent 'Q1' shear quality is commonly associated with high propagation propensity seen in skier-triggered avalanches.

Observed Fracture Type	**Fracture Character Code**	**Typical Shear Quality**
Sudden Collapse	SC	Q1
Sudden Planar	SP	Q1
Resistant Planar	RP	Q2
Progressive Compression	PC	Q3
Non-planar Break	BR	Q3

They argued that observing shear quality (or fracture character) in general could improve snowpack stability assessment through what they suggested was a "qualitative measure (at a small scale) of how well a fracture will propagate through a weak layer" (Johnson and Birkeland, 2002, pp. 5). Birkeland and Chabot (2006) later showed that interpreting shear quality in the rutschblock, stuffblock, and compression test could reduce the number of false-stable results (results indicating stability when in fact the slope is unstable) by four percentage points when compared to interpretation based solely on test score.

Johnson and Birkeland (2002) also revealed that their experience led them to believe that shear quality (fracture character) was less spatially variable than test score, which sparked numerous investigations on the subject (e.g. Landry, 2002; Kronholm, 2004; Campbell, 2004; Campbell and Jamieson, 2007; van Herwijnen et al., 2009) that arrived at similar general conclusions but generally lacked convincing evidence (for a review see Schweizer et al., 2008).

2.2.2 Rutschblock 'release type' and 'fracture quality'

In Switzerland, Schweizer and others (1995a) had developed a similar scheme for noting both 'fracture quality' and 'type of release' in the rutschblock test and Schweizer and Wiesinger (2001) and Schweizer (2002) had suggested integrating those observations into snowpack stability evaluations as a potential measure of propagation propensity. Fracture quality was described as 'clean', 'partly clean', or 'rough', and type of release was classified as 'whole block' 'most block' (i.e. below skis) and 'edge of block'. Since the dimensions of the rutschblock test likely exceed the critical failure length required for fracture propagation in most slab/weak layer combinations (Schweizer, 1999), it was reasonable to expect that the rutschblock results may provide some insight into the propagation propensity of that layer. Schweizer and Wiesinger (2001) proposed using rutschblock score, along with observations like layer hardness and fracture quality, to determine a stability rating, but noted that the rutschblock score criterion for each stability level is only fulfilled if the whole block released on a clean fracture plane. For example, an RB score of 2 or 3 indicated 'poor' stability but if the fracture was not clean or the whole block did not release, then higher stability could be considered due to expected lack of propagation. Schweizer and Jamieson (2003) showed that both release type and fracture type (character/quality) correlated with human triggering, and that release type was the more significant of the two. Jamieson (2003) argued that by including these observations, interpretation of RB scores could be improved. In studies of spatial variability in the rutschblock and other stability tests, Campbell (2004) and Campbell and Jamieson (2007) found that sudden fractures associated with higher propagation propensity occurred consistently throughout avalanche start zones and were less variable than point stability (test scores).

2.2.3 Further techniques for evaluating propagation propensity

The deep tap test (DTT) (introduced in Chapter 1), and its more quantitative (and now obsolete) cousin the fracture resistance test (FRT), both use the same column dimensions as the compression test but are applied to deeply buried weak layers by removing the upper column to within 15 cm of the targeted layer before applying surface load (Jamieson, 2003). Because the upper snowpack is removed, the DTT (or FRT) is not so much an index of skier-triggerability from the surface as it is a qualitative estimate of fracture propagation based on the observed fracture character (e.g. van Herwijnen and Jamieson, 2003). Presently, no research has been done to correlate fracture character in the DTT to deep-slab avalanches.

The merit of using other profile observations such as weak layer grain type and changes in hardness or grain size at layer boundaries for evaluating snowpacks potentially favourable to propagation have been reviewed (e.g. Schweizer and Jamieson, 2003; McCammon and Schweizer, 2002; Schweizer et al., 2004; Jamieson and Schweizer, 2005; van Herwijnen, 2005; van Herwijnen and Jamieson, 2007b; Schweizer and Jamieson, 2007). The snowpack characteristics strongly associated with snowpack instability are commonly called ‘Lemons’ (McCammon and Schweizer, 2002) or ‘Yellow Flags’ (Jamieson and Schweizer, 2005) and may be better indicators of fracture initiation *and* propagation potential than the initiation scores of common stability tests like the compression test and rutschblock (Jamieson and Schweizer, 2005).

More recently, Schweizer and others (2006) evaluated combining rutschblock score, fracture quality and release type with an index of structural stability (lemon/yellow flag count) to estimate skier triggering probability in terms of fracture propagation propensity. Results showed that rutschblock release type was the best single predictor (71% accuracy) but that all three used in combination improved the stability assessment. Schweizer and Jamieson (2007) expanded on the model of a structural stability index to establish critical values for the Yellow Flags that could be combined in a 'threshold sum' to accurately discriminate between stable and unstable slopes, however they did not incorporate fracture type (character/quality) in stability tests for that analysis.

2.2.4 Summary: a need to specifically test propagation potential

In their review of snow avalanche formation, Schweizer and others (2003a) pointed out that all traditional stability tests try to reproduce, to some degree, the dynamic surface loading of a skier (snowboarder, snowmobiler etc.) that causes brittle fracture in weak layers, but that none provides direct information on fracture propagation propensity. Instead, they primarily test the ease of fracture initiation through a qualitative measure of weak layer strength. Much earlier, Jamieson and Johnston (1992) had noted that the width of slab avalanches (propagation distance) appeared independent of the initial trigger energy, which all traditional stability tests measure in some way. Schweizer and others (2003a, 2006) also pointed out the common occurrence of stability tests indicating instability or weakness when no propagation is observed in the field (false-unstable), inferring that traditional stability tests miss important information about propagation potential and hence the probability of slab release.

In a recent comparison of the prediction accuracy of popular stability tests on skier tested slopes (both stable and unstable), Gauthier and Jamieson (2008b) showed that overall prediction accuracy ranged between 69% for RB score and 79% for CT fracture character. False-stable results in these tests ranged from 7% for CT fracture character and Yellow Flags to 20% and 26% for RB release type and RB score, respectively. Yellow Flags and CT fracture character showed higher rates of false-unstable results (19% and 16% respectively) which reinforces the observation of Schweizer and others (2003a) that non-propagation-specific stability tests tend to overestimate instability.

Despite the improved interpretation of common stability tests when fracture character (shear quality/release type etc.) is included in addition to test score (e.g. Birkeland and Chabot, 2006, van Herwijnen and Jamieson, 2007a; Schweizer et al., 2006; Gauthier and Jamieson, 2008b), false-stable results still constitute a significant percentage of test results, which can prove fatal to the operator or observer. Also, it was commonly noted in spatial variability studies using stability tests (e.g. Kronholm, 2004; Campbell, 2004, Campbell and Jamieson, 2007) that even if no fracture initiated in a deep weak layer within a stability test, avalanches could release by fracture initiation and subsequent propagation on the same weak layer in a shallower snowpack area nearby. This also partially accounts for the difficulty in predicting destructive deep-slab avalanches, since other that the deep tap test, no other snowpack test provides information on deeply buried weak layers beyond the surface-triggerable range.

The information gaps surrounding fracture propagation, the imperfect stability prediction accuracies, and the dependence on surface loading in existing column stability tests raises numerous questions about the validity of these methods for assessing propagation propensity: could the prediction accuracy of standard stability tests be bettered if snowpack tests are used that specifically target and index the propagation propensity of the weak layer and slab combination? – Perhaps independently of the ease of fracture initiation via surface loading? Learning from recent theoretical developments on fracture propagation, could a test be developed and perfected to mimic the true physical propagation phenomenon beyond the point of fracture initiation, as observed in natural and artificially triggered avalanches and whumpfs? The following sections review recent efforts to answer these questions and provide the framework for the further development of the Propagation Saw Test.

2.3 Development of the Propagation Saw Test (PST)

2.3.1 Early prototype propagation tests

Gauthier and Jamieson (2006a, 2006b) first introduced prototype field tests for specifically assessing fracture propagation in weak layers using an elongated column after the 2005 winter; however, the origin of the PST can be dated back to when B.C. Johnson (2000) used the cantilevered beam test to study whumpfs and remotely triggered avalanches. The cantilever beam test effectively measures the flexural strength of snow (Perla, 1969) and is not necessarily performed on a weak layer, but commonly was in Johnson's work. It shared a similar method and dimensions to the modern PST, but differed in that the back (up-slope) end was not isolated from the snowpack (forming the cantilever), and that a thicker saw was used to undercut the beam, effectively creating the cantilever. In his research, B.C. Johnson (2000) found that when no weak layer was present, a vertical slab fracture developed through the beam typically within 5 cm from the end of the undercut, but that when a weak layer was undercut, the slab fracture frequently developed 30-60 cm from the end of the cut, and often at the end of the column. He observed that the fracture initially propagated from the end of the saw cut and was interrupted by the subsequent slab fracture.

Gauthier and Jamieson's (2006a) first prototype test used a longer (3 m) column which was fully isolated from the surrounding snowpack, and used a drop-hammer apparatus (e.g. Stewart, 2002) to initiate a fracture which could then propagate in the column. They justified the longer column by noting that traditional small column stability tests like the compression test, stuffblock, or drop-hammer test likely had insufficient area to observe propagating fractures independently of the fracture's initiation; and that a suitable propagation test should allow the development of a flexural wave in the slab as seen in recent field experiments (B.C. Johnson, 2000; van Herwijnen and Jamieson, 2005) and described theoretically by Heierli and Zaiser (2006, 2007). Gauthier and Jamieson expected that any observed propagation in the column would be independent of the initiation energy (drop-height) as had been observed in skier-triggered avalanches by Jamieson and Johnston (1992).

The same winter, Gauthier and Jamieson (2006b) experimented with the now-standard method of drawing a 2 mm thick saw blade through the weak layer until the onset of propagation, although they originally started from the up-slope end. They tested columns ranging from 30 cm to 4 m in length and found that at some cut length in all 101 tests a fracture propagated ahead of the saw either to the end of the column or until it stopped at a fracture through the overlying slab. This led them to suggest that a relationship between cut length and isolated column length might exist where either a constant linear proportion of the column had to be cut to initiate fracture propagation; or where a constant, absolute cut length was required regardless of column size, but that insufficient data existed to confirm either. Either way, they were convinced that shorter cut lengths would be required when propagation propensity was high, regardless of the column length.

Concurrently and independently from Gauthier and Jamieson, Sigrist (2006) and Sigrist and Schweizer (2007) developed a nearly identical test method that they used in combination with a finite-element model for determining the critical energy release rate required for fracture propagation within weak layers. They used column lengths of 0.6 m, 1.2 m, and 1.8 m and cut through the weak layer from either end in a smooth sawing motion.

2.3.2 The "Propagation Saw Test"

After another winter of data collection, Gauthier and Jamieson (2007a) and Gauthier (2007) examined the relationship of cut length and column length, and many other test variables, to establish the current test method and column dimensions, which was eventually coined the Propagation Saw Test, or PST (Gauthier and Jamieson, 2007b). The controlled variables of column length, slope angle and cut direction were altered through more than 600 tests and compared independently in groups of tests from the same day and location (or arguably similar snowpacks on subsequent days) by isolating one variable for the comparison. Simultaneously, natural variables of the snowpack including weak layer depth, weak layer crystal type, and slab density differed between test groups which enabled assessment of their effect on test results. They used the Mann-Whitney *U*-test (Mann and Whitney, 1947) to ensure that any observed lack of difference between compared variables was statistically significant. Their results indicated that cut direction and slope angle had no significant influence on the test results in most comparisons and slightly shorter cut lengths and longer propagation lengths were only observed occasionally when cutting in the up-slope direction and in tests on the flats. In terms of column length, they suggested that 1 m or equivalent to the slab depth would be long enough to gain sufficient information about the critical cut length required to initiate fracture propagation and about the subsequent propagation length that followed.

Gauthier and Jamieson abandoned the earlier drop-hammer method in favour of the saw-cut method since it effectively separated the fracture propagation process from surface-induced fracture initiation, and because while the traditional surface-loading tests were limited to layers around 1 m deep in the snowpack, the PST enabled the testing of deeply buried weak layers – up to 3 m or more (Gauthier and Jamieson, 2007b). Gauthier and Jamieson (2006b, 2007a), Gauthier (2007) and Sigrist and Schweizer (2007) described the saw-cut initiation method as being consistent with both the shear fracture mechanics models and weak layer collapse models presented in Section 2.1.

2.3.2 Validating the Propagation Saw Test

Early validation of PST results next to whumpf and avalanche sites (unstable sites) or at sites where propagation was *not* observed on a skier-tested slope (stable sites) solidified the current test method and geometry, and generated the simple interpretation rule

(Section 1.8.2) presently used for predicting propagation propensity on the slope scale (Gauthier, 2007; Gauthier and Jamieson, 2007b, 2008a). Based on an initial dataset of 23 unstable and stable slopes, Gauthier and Jamieson (2007b) showed that the proposed interpretation rule accurately predicted avalanches and whumpfs in most cases but did show a high number of false-stable predictions, particularly for thin, soft slabs. Subsequent validation efforts (Gauthier and Jamieson, 2008a, 2008b; Gauthier et al, 2008) showed similar prediction accuracies and attributed false-stable results to thin, soft slabs, large weak layer crystals, and longer columns used prior to the current standard length. Table 2.2 summarizes the different methods and results of these initial validation studies.

Table 2.2: Previous validations studies on the PST using various methods and datasets.

Validation Study	**Sites/Tests**	**% Correct**	**% False-Stable**	**% False-Unstable**	**Method Differences**
Gauthier and Jamieson, 2007b; 2008a	23 / 170	75%	20%	5%	Included all column lengths from 0.2 to 2.6m at stable and unstable sites
Gauthier and Jamieson, 2008b	27 / 187	77%	18%	5%	Included all column lengths as above and skier-tested sites without confirmed initiation
Gauthier et al., 2008	33 / 67	73%	27%	0%	Only included standard column lengths and only avalanche/whumpf (unstable) sites.

Standard and modified versions of the PST column and test method have been used to develop and confirm some of the theory described in Section 2.1, particularly the long columns used by van Herwijnen et al. (2008) for monitoring propagating fractures with a high-speed camera; the columns of variable slab thickness used by Simenhois and Birkeland, (2008a); and the fixed-back columns used by McClung (2009a) to support the ratio of critical cut length to slab thickness used in the shear fracture propagation model (e.g. Bažant et al., 2003; McClung, 2005a).

2.4 Development of the Extended Column Test (ECT)

The ECT was developed concurrently but independently of the PST as a test for fracture initiation *and* propagation and was first presented by Simenhois and Birkeland (2006) after an initial season of testing in Colorado and New Zealand. It employs the same loading method as the compression test (Section 1.7.1) but uses an elongated, 90 cm across-slope column and requires a unique recording method to evaluate both initiation and propagation (see Section 1.8.1). Their results from 256 tests on stable slopes and 68 tests on unstable slopes showed that the ECT accurately predicted unstable slopes 100% of the time based on the rule that propagation was *likely* only if the fracture propagated across the column within two additional taps of fracture initiation, and it accurately predicted stable slopes 98% of the time with only 2% of tests falsely predicting instability on a stable slope (Simenhois and Birkeland, 2006). Subsequently Simenhois and Birkeland (2007; Birkeland and Simenhois, 2008) refined their interpretation rule to allow only one additional tap after initiation within which propagation can still be rated *likely*; and presented the current recording standard (Section 3.4). Additional validation data were also presented (Simenhois and Birkeland, 2007, 2009; Birkeland and Simenhois, 2008) with comparably high correct prediction rates (> 90%) and only 3-4% false-stable predictions on unstable slopes. It is important to note that much of the data used to validate the ECT in these subsequent studies (Simenhois and Birkeland, 2007; Birkeland and Simenhois, 2008) came from the SnowPilot (Chabot et al., 2004) database in which stable and unstable slopes are determined based on the user's stability rating at the time (i.e. poor, very poor stability rating = unstable slope; good, very good stability rating = stable slope), which was often influenced by the ECT or other stability test results. This differed from the more conventional method of skier testing the slope to determine stable versus unstable and may have generated a higher number of accurate predictions since the slope rating to which the test was compared may have been dependent on the test itself.

The effectiveness of the ECT in predicting stability compared to more traditional methods (RB, CT, structural stability indices) was more recently tested in the Swiss Alps (Winkler and Schweizer, 2008, 2009) and the Eastern Pyrenees (Moner et al., 2008) and comparable results were reported. Simenhois and Birkeland (2006, 2007) also report on two sets of spatial arrays in which the ECT showed spatial uniformity on a stable slope (2006) and reflected the variability in conditions on a second slope (2007).

It is important to note that these studies all use the ECT to predict stability in terms of the probability that a skier will trigger an avalanche on a slope. This is arguably different than the probability that a skier will trigger a failure within a weak layer (CT score) and does not effectively predict propagation propensity in cases where the test does not initiate fractures in deeper layers, as observed by Simenhois and Birkeland (2009).

The ECT, as well as the PST, has additionally been used to study other phenomena effecting propagation. Exner (in preparation) is using the PST to study the effect of daytime warming on propagation potential and other snowpack characteristics, while Simenhois and Birkeland (2008b) used the ECT in multiple case studies to conclude that surface warming can increase propagation potential in buried, dry weak layers. Simenhois and Birkeland (2008a) used both the PST and ECT to study the effect of changing slab thickness within a test column and determine that fractures typically propagated more readily from thin to thick areas. Hendrikx and others (2009) used the ECT to assess the spatial variability of fracture propagation propensity over time and found that results could vary significantly in both time *and* space over the same slope.

2.5 Snowpack Characteristics Conducive to Fracture Propagation

In the development of snowpack tests and their application towards understanding fracture initiation and propagation in avalanche release, numerous studies have reported on the slab and weak layer properties and combined characteristics that appear to strongly affect propagation propensity. For fractures to self-propagate in a weak layer, Sigrist and Schweizer (2007) generalized that sufficient energy must be released in the failure process that depends on the material properties of the slab and the collapsible height of the weak layer. Schweizer and Componovo (2001) more specifically suggested that propagation would be dependent on stiffness differences between the weak layer and the slab, specifically the slab layer immediately above the weak layer. While studying fracture character in compression tests, van Herwijnen and Jamieson (2004a) also hypothesized that the characteristics of the layer immediately overlying the weak layer contributed to propagation propensity. They later confirmed this hypothesis (van Herwijnen and Jamieson, 2007b) and additionally found that pronounced hardness differences and crystal size differences across the failure plane favoured fracture initiation and propagation, while stiffer overall slabs favoured propagation. This latter result compared well to observations of large propagation initiated from low angle terrain in remotely triggered avalanches and whumpfs (B.C. Johnson, 2000), that tended to correlate with thicker, denser, and harder slabs.

For the development of structural stability indices (Yellow Flags, Lemons) McCammon and Schweizer (2002), Schweizer and Jamieson (2003), Jamieson and Schweizer (2005), and Schweizer and others (2004) evaluated snow-profiles next to skier-triggered slopes and stable slopes and determined that weak layer depth, large differences in grain size and hardness across the failure interface (hardness transitions), and large and persistent grains in a soft failure layer were indicative of snow instability both in terms of fracture initiation and propagation. More recently, Habermann and others (2008) used a finite-element model to show that stiffer layers *below* the weak layer increase stress concentrations in the weak layer favouring fracture initiation. They also concluded that while stiffer overlying layers may increase propagation propensity (e.g. van Herwijnen and Jamieson, 2007b), they may also reduce surface induced stress in the weak layer due to a 'bridging' effect that makes fracture initiation more difficult.

2.6 Forecasting for Deep-Slab Avalanches

Few studies have attempted to improve our understanding of deep-slab avalanches or suggest techniques and observations that may help forecast such avalanches. The term deep-slab is synonymous with 'a thick slab overlying a deeply buried weak layer'. Jamieson and others (1998) showed that accumulating snowfall or large temperature changes over multiple days correlated with natural deep-slab avalanches (average crown thickness > 85 cm) that released on a particular facet-crust layer in 1996-97, but argued that these effects were coupled with starting zones where slab thickness was variable with local thin spots. They also included the shear-frame stability index in their observations and found a weak correlation with deep-slab avalanche activity within a study area of 100 km around the test site. Schweizer and others (2009) described similar factors producing infrequent deep-slab avalanches, but particularly focused on critical amounts of accumulated new snow. No current research has specifically used a snowpack test to help forecast deep-slab avalanches.

Although the PST would not be able to predict *when* deep slabs may release, it may be able to indicate whether weak layers beneath deep slabs possess the *potential* for propagation. This could help practitioners and recreationists *anticipate* deep-slab avalanches during large snowstorms or temperature fluctuations, or where shallow snowpack spots may facilitate artificial triggering.

2.7 Regional Validation of Snowpack Test Results

Practitioners have often tried to extrapolate point source information from profiles and stability tests to create forecasts of avalanche conditions over large areas of terrain. By similar means, researchers have attempted to validate test methods against stability or avalanche observations on a regional scale, often tens of kilometres away from the test site. In a study of regional forecasting, Jamieson and others (2008) showed that local "nowcasts" derived by experienced local observers matched the danger ratings reported in the public avalanche bulletin for the surrounding region in approximately 59% to 64% of cases across Western Canada. In a similar study that employed profiles and stability tests as opposed to "heads-up" observations of local surroundings, Schweizer and others (2003b) found that local conditions often varied greatly from the prevailing regional forecast. Both these studies exemplify the challenge of regional forecasting, and the potential danger of verifying regional forecasts based on single test results or limited local observation. Despite this, Hägeli and McClung (2003) showed that large-scale weather events, including the formation of PWLs, were fairly consistent over the entire Columbia Mountains. This is discussed further in Section 3.8.1, but may reveal that, although variability of some conditions may be high from one specific area to the next, significant weather and snowpack conditions can show consistency over huge areas.

A few studies have compared specific test results to regional avalanche activity. Jamieson and Johnston (1993b) showed that stability parameters calculated from shear-frame test results correctly predicted some potentially harmful (size 1.5+) regional avalanching within a 30 km radius on at least 75% of the days in which they were evaluated. Using the same approach, Jamieson (1995, pp. 180) showed rutschblock scores ≤4 frequently corresponded with regional (~15 km radius) skier-triggered avalanche activity on the same PWL within a day of the test, and that, in general, regional skier-triggered avalanche activity decreased with increasing rutschblock scores at study sites. More recently, Jamieson and others (2007) revisited the use of stability indices in regional forecasting and concluded that spatial variability between study plots and across a region can greatly influence the success of stability index predictions extrapolated from tests at a study plot; and that regional avalanche activity likely correlates better with the stability index *trend* rather than the specific value on a given day.

Despite the mixed success of comparing avalanche activity on the regional scale to local (site-specific) test results and stability observations, the approach *may* prove valuable in validating PST results (particularly multiple-day trends in propagation), since it has been widely argued that propagation propensity is less spatially variable across the slope and regional scale than other conditions such as triggerability (initiation ease) (e.g. Johnson and Birkeland, 2002; Campbell and Jamieson, 2007; van Herwijnen et al., 2009).

2.8 Summary

The literature reviewed in this chapter introduced the key theoretical developments that describe fracture propagation in weak layers as part of the avalanche release process. Much of this theory prompted the inclusion of 'propagation potential' observations (i.e. fracture character, shear quality, release type) in traditional snowpack tests, and inspired the development of new, propagation-specific snowpack tests. Some of the theory is the direct result of experiments with such tests in the field. Through either scenario, fracture propagation potential has proven to be of vital importance to the avalanche release process and of great interest to forecasters, guides, and experienced winter recreationists who wish to be able to predict its propensity. While inclusion of 'propagation potential' observations in traditional stability tests like the CT and RB and in snow-profiles (Lemons/Yellow Flags) have proven valuable, the development of the ECT, and particularly the PST, give experienced practitioners – and perhaps eventually recreationists – additional tools to specifically test the propagation propensity of concerning weak layers.

Recent comparisons of test results (Birkeland and Chabot, 2006; Gauthier and Jamieson, 2008b; Schweizer and Jamieson, 2009; Moner et al., 2008; Winkler and Schweizer 2008, 2009; Simenhois and Birkeland, 2009) reveal that no test is perfect at predicting instability in all cases at all times, but that the newly developed ECT and PST appear to be better suited to accurately assess propagation propensity without overestimating instability, as evident by their low number of false-unstable predictions (e.g. Simenhois and Birkeland, 2009; Gauthier and Jamieson, 2008b). However, in the small validated dataset that existed prior to this study, the PST has been shown to give more false-stable predictions than other tests, including the ECT (Schweizer and Jamieson, 2009; Simenhois and Birkeland, 2009). A purpose of this study is to develop and expand on the existing validation dataset to determine how accurately the PST's predictions represent true stability on surrounding slopes, and to compare the new results with other snowpack tests, particularly the ECT.

The PST's independence on surface loading allows for the evaluation of propagation propensity in deeply buried weak layers which before only the deep tap test attempted. This could potentially help forecasters anticipate deep-slab avalanches, which this study attempts to validate along with improved predictions in shallow, soft slabs. Furthermore, validation on a regional scale may reveal whether propagation propensity truly is less variable than triggerability, and may establish the PST as a suitable tool for regional avalanche forecasting.

3. METHODS

3.1 Study Areas and Field Sites

Fieldwork for this study was conducted entirely in the Columbia Mountains of interior British Columbia (Figure 3.1). The Columbia Mountains have a transitional snow climate (Hägeli and McClung, 2003) characterized by a mixture of heavy maritime-like snowfall, with periods of clear weather that can produce persistent weak layers (PWLs), primarily surface hoar. Melt-freeze crusts commonly form early-season, and are often accompanied by near-crust faceting. Annual snowfall regularly exceeds ten metres and temperatures are generally moderate, with warmer average temperatures than the Rocky Mountains to the east of the Columbia Valley, and a slightly drier snow climate than the Coast Range to the west.

The Applied Snow and Avalanche Research group at the University of Calgary (ASARC) operated two permanent field stations in the Columbia Mountains for the duration of this study: one based out of Rogers Pass in Glacier National Park between Golden and Revelstoke; and a second at Mike Wiegele Helicopter Skiing in Blue River. Additional fieldwork was occasionally conducted outside the boundaries of Kicking Horse Mountain Resort near Golden, BC, and at Chatter Creek Cat Skiing north of Donald, BC.

Two regular study areas were used in Glacier National Park and within the tenure of Mike Wiegele Heli Skiing. Mount Fidelity is at the west end of Glacier National Park and has been the site of the highway Avalanche Control Section's (ACS) regular study plot and weather observation for many years. In addition to a snowpack that is representative of park conditions up to and just east of the summit of Rogers Pass, Fidelity provides an excellent winter-long study area with its protected low-angle slopes and easy access via a snow-road. The same applies to Mount Saint Anne in the Cariboo Mountains northwest of

Blue River where representative weather and snowpack conditions and snow-road access make it ideal for a regular study area within the Mike Wiegele tenure. Mount Abbott hosts another representative study site in Glacier National Park with more difficult access and thus less-frequent use. All three study areas are around 1900 m above sea level.

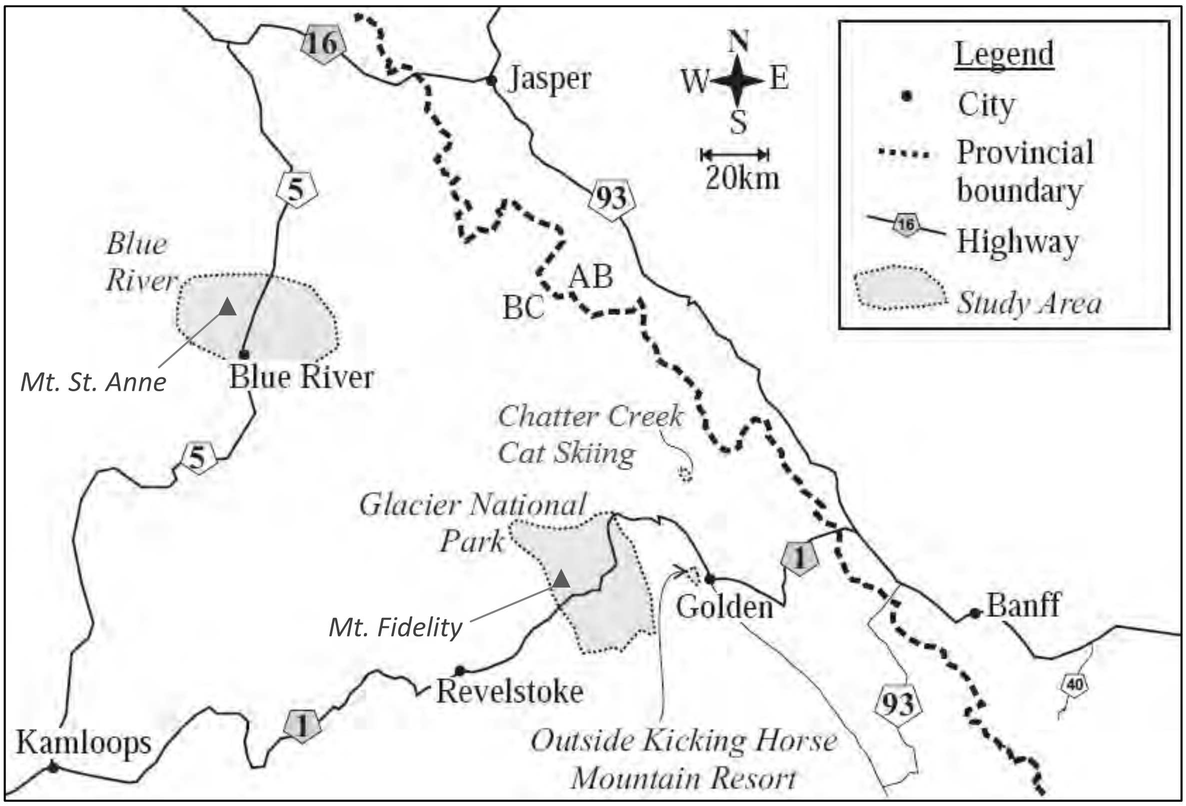

Figure 3.1: A map of the Columbia Mountains of interior British Columbia, showing the location of Glacier National Park and the Mike Wiegele Helicopter-Skiing tenure around Blue River, BC, where the majority of fieldwork was conducted. Kicking Horse Mountain Resort and Chatter Creek Cat-Skiing are also shown on the map (after Gauthier, 2007).

Although the two study areas of Mt. Fidelity and Mt. Saint Anne were used frequently, most of the data for this study were collected at a variety of sites and locations throughout Glacier National Park and around Blue River. These sites were termed 'roaming sites' and included any location not regularly used as a study site. V*alidation* sites, which are described in detail in Section 3.5, were sites of confirmed propagation (avalanches or whumpfs) or confirmed fracture initiation *without* propagation (unreleased ski-cuts) used to validate the PST. They were often on skier-tested slopes in the study areas of Mt. Fidelity or Mt. St. Anne, but also consisted of numerous roaming sites where potentially skier-triggerable slabs were sought out, or where an avalanche had occurred over the previous 24 hours. Most sites were at or below tree line elevation and covered the full range of aspect and slope angles from zero degrees to over fifty degrees. Selected slopes had to have sufficient room and uniformity to perform the required fieldwork (Section 3.2) but also had to be small enough or low-angle enough not to present an avalanche hazard to field workers. Often, known PWLs or other layers of interest were targeted for testing and included surface hoar layers, faceted layers around or within melt-free crusts, and a few shallow storm interfaces and wind-slabs. Throughout the two seasons, the tested layers ranged in depth from 10 cm to 250 cm below the surface.

3.2 Fieldwork

Fieldwork for this study was conducted in the winters of 2007-08 and 2008-09. Although some data were collected in late December of each winter, most were collected in January to March of the new year and thus the field seasons are henceforth referred to as the 2008 and 2009 winters.

Daily fieldwork differed slightly between the winter of 2008 and that of 2009 based on the developing objectives of this study. On all field days, a new snowpit was dug and at least two standard PSTs (Section 3.3) were performed in the snowpit along with two compression tests (CT) and a snow-profile as described in Chapter 1 and in conformance with the CAA Observation Guidelines and Recording Standards (OGRS) (CAA, 2007). The profile was conducted to at least one layer below the layer of interest and included temperature, grain size and type, hand-hardness and densities for each layer as described in OGRS. In addition to the standard PSTs, CTs and profile, two ECTs were performed in each pit in 2008 (Section 3.4), while two short-scaled PSTs (Section 3.6) replaced the ECTs in 2009. In both seasons, any opportunity to test the PST at validation sites of confirmed propagation or confirmed initiation without propagation was exercised. Frequently, more than two standard PSTs were performed in a snowpit. The occasional rutschblock (RB) or deep tap test (DTT) was also performed in conformance with OGRS. Figure 3.2 shows the typical layout of snowpits during the 2008 (a) and 2009 (b) field seasons. Between multiple tests performed in a single snowpit, sufficient snow surrounding each test (15 cm minimum) was cleared away to ensure the slab and weak layer were entirely intact for the subsequent test.

Figure 3.2: Typical snowpit layouts from the winters of 2008 (a), and 2009 (b). A minimum of 15 cm of snow was cleared away between test columns to ensure the slab and weak layer were intact in each subsequent test.

Site selection required sufficient space to accommodate the profile and required tests, plus room for additional tests if initial tests were inconclusive or if operator error was deemed to be a factor (e.g. saw cut exiting the weak layer in a PST). To reduce error and ensure consistency, the same operator performed every test of the same type (e.g. CT, PST etc.) on a given field day. Any tests obviously subjected to operator error were discarded and repeated.

Measurements of slope angle (Ψ), total snowpack depth (*HS*), aspect (cardinal direction or degrees), air temperature (^{o}C), elevation (metres or feet), and geographic coordinates were recorded at the site, along with a description of the site location and any notes regarding alternate methods, avalanche observations, or ski-cut outcome.

The author was not present for the collection of all data used in this study. The author *was* present approximately 73 of 98 (75%) PST days at Rogers Pass and 8 of 89 (9%) of PST days at Blue River. Quality control was ensured through careful methods training with all ASARC students and technicians at the beginning of each season, and through daily conference calls between field stations.

3.2.1 Winter 2008

In the 2008 winter, 91 snowpits on 88 field days were devoted to comparing the PST and ECT, many of which coincided with the validation study at sites of confirmed propagation or initiation without propagation. An additional 10 pits were devoted to testing deeply buried weak layers with the PST only. Further PSTs were performed throughout the winter as parts of other ongoing ASARC studies. Although numerous different types and individual weak layers were tested throughout the season, the majority of data in 2008 was collected on four PWLs that were identified (ID) by burial date as the 5 December (2007) rain crust, the 26 January surface hoar, the 23 February surface hoar, and the 9 March surface hoar. All four layers had widespread prevalence throughout the Columbia Mountains and other areas of BC and Alberta, often with burial dates (IDs) that varied slightly. The existence of two reactive PWLs close to each other in the snowpack late in the season allowed for multiple results – and hence multiple comparisons – from within the same PST and ECT columns. In other words, two layers could be cut in one PST column and would both usually react in the ECT. It was determined that multiple layers could be tested in a single PST column provided the operator worked from the bottom layer up, maintaining the integrity of the overlying slab during each test.

3.2.2 Winter 2009

In the 2009 winter, the standard PST was performed in 101 pits on 89 field days. A few of those days were intended for other ASARC research projects but the majority of them were devoted to comparing short-scaled PSTs next to standard PSTs, and to validating both standard and short-scaled tests next to sites of confirmed propagation or initiation without propagation. The short-scaled PST involved scaling the column length to match weak layer depths *less than* 1 m, as opposed to only scaling the length beyond 1 m as in the standard PST. The short-scaled PST method is described in detail in section 3.6.

As in 2008, various layers were tested, but four PWLs dominate the dataset of 2009. As identified by burial date, they were the 25 December (2007) facets, the 27 January surface hoar, the 22 February surface hoar, and the 1 March surface hoar. All three surface hoar layers in 2009 were commonly found above a thin melt-freeze crust and associated facets that formed during the same clear-weather period on south to west aspects. Again, all four layers were common throughout the Columbia Mountains and large parts of BC and Alberta with slightly different burial-date IDs.

3.2.3 Field equipment

Equipment for this study consisted of a simple snow-observation kit and a snow-saw carried by most practitioners and some recreationists, as well as a collapsible 240 cm+ probe and a snow-shovel ideally carried by anyone in avalanche terrain. The saw used by all ASARC field workers for every PST was a 45 cm long, 5 cm wide, and 2 mm thick LifeLink® snow-saw. The snow-observation kit contained a foldable ruler for measuring layer depths in the profile and column dimensions and results in snowpack tests; a digital thermometer for accurately measuring air temperature and snowpack temperatures; a magnifying loupe and crystal screen for observing grain type and size in the profile; and a long (typ. 10 m) piece of 3 mm diameter cord used in combination with probes for isolating test columns via cord-cutting. In addition, a density kit consisting of a drug-scale and a small 100 cm^3 sample tube was used to measure layer densities in the profile.

3.3 Standard Propagation Saw Test Method

The standard PST introduced in Chapter 1 and performed throughout the 2008 and 2009 winters uses a column that is 30 cm wide across slope by 100+ cm long up-slope, isolated on all four sides from the surrounding snowpack to below the targeted weak layer as originally proposed by Gauthier (2007 pp. 197). The column was typically isolated at the front and on one side completely by shovel and at the back and remaining side by cord-cut or occasionally by vertical saw-cut (Figure 3.3a). The length of the column was always 100 cm (+/- 5% typically) when testing layers less than 100 cm deep, but was scaled to match layer depth when the layer was deeper than 100 cm (Gauthier, 2007 pp. 197). For example, a layer buried 65 cm deep would be tested with a column 100 cm long, while a layer buried 132 cm deep would be tested with a 132 cm long column.

Once isolated, the weak layer was carefully identified across the front and length of the column, often using a small, soft brush or by brushing with the back of a glove to ensure it was easy to follow visually. Fracture initiation was then simulated by steadily (~15-20 cm/s) drawing the blunt edge of the saw up-slope within the weak layer (saw-cutting), carefully as not to draw the saw out of the layer (Figure 3.3b). This process of artificial fracture initiation was continued (Figure 3.3c) until the fracture suddenly propagated ahead of the saw, or until the entire length of the column had been cut.

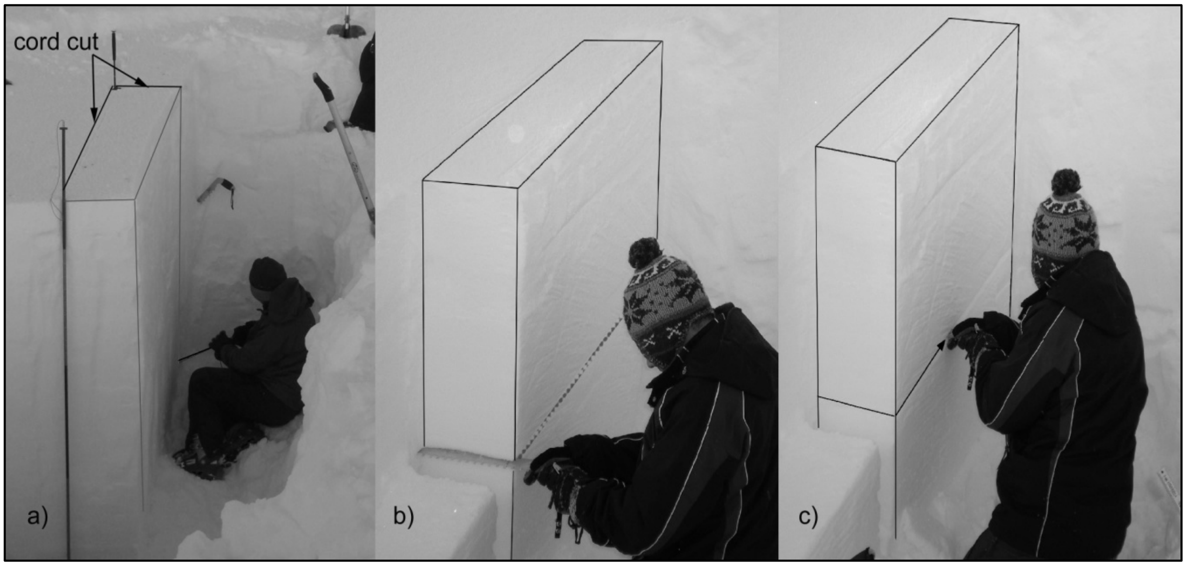

Figure 3.3: The standard PST column is isolated from the surrounding snowpack via cord-cut to below the weak layer of interest (a), after which the weak layer is identified across the column (b) and saw-cut up-slope (c) until the fracture propagates ahead of the saw.

With the onset of propagation, three different results are possible (Figure 3.4). Either the fracture propagates along the entire length of the column to the end (*END*), or the propagating fracture stops within the column at a slope-normal fracture through the overlying slab (*SF*), or at a point of self-arrest along the layer (*ARR*) (Gauthier and Jamieson, 2007a). When the operator observed even the slightest propagation ahead of the saw, he or she immediately ceased cutting and marked the point from which propagation started either with a finger on their free hand, or by turning the saw blade down into the layer below.

Measurements of isolated column length (y), depth to weak layer (z), cut length (x), and propagation length (l_p) were taken and recorded in a field book along with the observed propagation character (*ARR/SF/END*), the tested weak layer date ID (*yymmdd*), and any other comments regarding the test method or result.

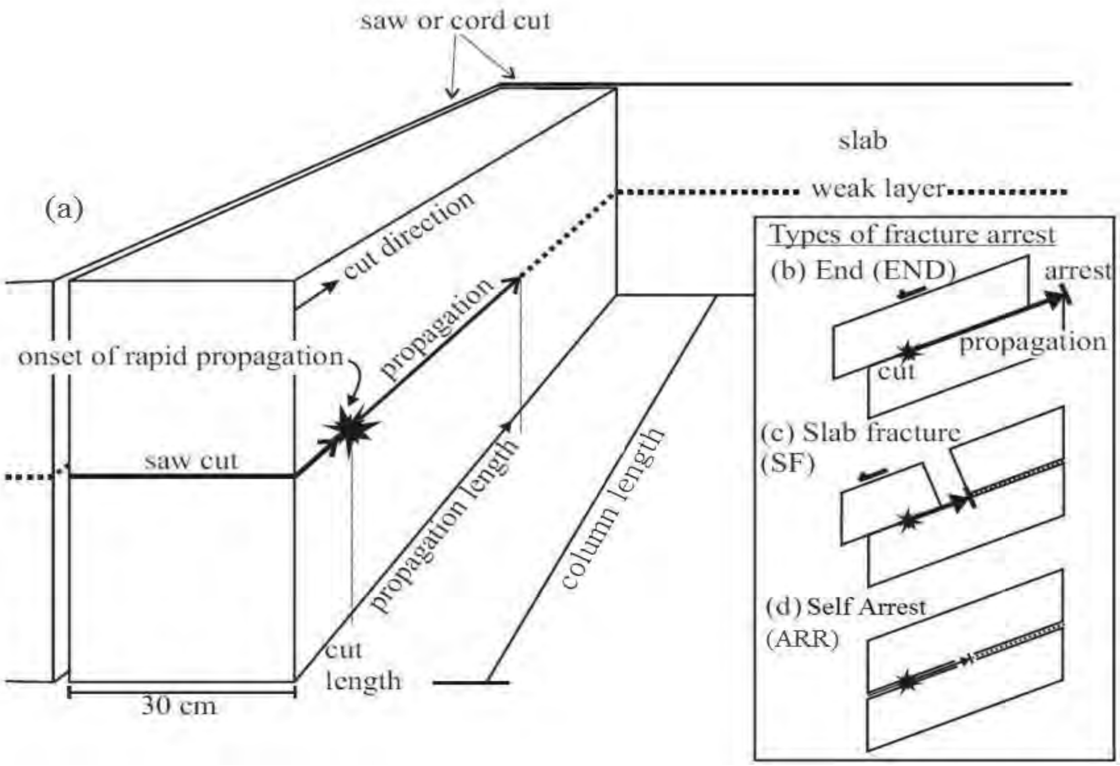

Figure 3.4: A schematic of the standard PST method (a) and the three possible results of END, SF, and ARR obtained in the test (b-d) (after Gauthier, 2007).

The standard recording method as originally proposed by Ross and Jamieson (2008) and revised slightly by Ross and Jamieson (2009) is ***x/y* (*ARR/SF/END*) down *z* on *yymmdd*** (or alternative layer ID) **on a *Ψ°* slope,** where *x* is cut length, *y* is isolated length, and *z* is weak layer depth. Slope angle Ψ is included in the recording standard due to a small dataset (McClung, 2008, 2009) contradicting the slope-independence of test results published by Gauthier and Jamieson (2007a) and Gauthier (2007), which is not resolved in this work. This was the recording method used to present our results publicly, and suggested to practitioners interested in using the test.

The PST results were interpreted as indicating propagation likely (*propL*) on surrounding slopes only when *less* than half the column (< 50%) had been cut at the onset of propagation, and the fracture propagated uninterrupted to the end of the column (*END*). If more than half the column had been cut when the fracture propagated to the end, or if the propagating fracture stopped at slope-normal fracture through the slab (*SF*) or a point of self-arrest (*ARR*) along the weak layer, propagation on surrounding slopes was said to be unlikely (*propUL*). This was as originally determined heuristically by Gauthier (2007 pp 127), and Gauthier and Jamieson (2007b) from early validation results.

Potential operator error was carefully monitored, particularly while performing the PST saw-cut. While isolating the column, longer columns were usually trimmed down to meet the standard length, although short or long columns were accepted if they were within ten to fifteen percent of the recommended standard length (e.g. 85 cm – 115 cm). Minor deviations from within the weak layer while drawing the saw through the layer were accepted, although any major deviations for more than a few centimetres of saw-cut were rejected and the test was repeated. In most cases, this was made obvious by the sudden resistance change while drawing the saw through the layer, but in cases of hard weak layers where dragging the saw was difficult, or in soft slabs over soft weak layers, deviations may have gone unnoticed.

Another potential error source was the measurement of cut length which the operator had to identify the moment fracture propagation began from the edge of the saw. Although this was usually straightforward, on some steep slopes when propagation was extremely rapid and the slab slid down-slope quickly, holding the saw still or maintaining a visual on the point at which propagation started was difficult. Dipping the saw into the substratum or marking the spot with a finger on the free hand helped in these cases. Additionally, after the test, the undisturbed wall behind the test was examined to see that the saw had scored the weak layer and thus penetrated through the column width for the entire length of the test. To reduce errors and improve consistency, one operator performed all PSTs on a given field day.

3.4 Extended Column Test Method

The Extended Column Test (ECT) was only performed as part of this study in the 2008 winter and followed the method originally proposed by Simenhois and Birkeland (2006). An `extended` CT column 30 cm up-slope by 90 cm across slope was isolated from the surrounding snowpack by shovel and cord-cut or vertical saw-cut (Figure 3.5a). The isolated test depth was often to below weak layers of interest and ranged from 45 cm to 190 cm, with an average test depth of approximately 110 cm. The blade of a shovel was then placed on the snow surface over one end of the column (Figure 3.5b), and the column was loaded in three stages of increasing force in the same manner as the CT: 10 `taps` from the wrist, 10 from the elbow, and finally 10 from the shoulder. The operator or an extra observer would watch for a fracture to initiate under the shovel in a weak layer within the column, and potentially propagate through the weak layer across the remainder of the column. In any such instance, the number of taps required to initiate failure and the subsequent number of taps to propagate the fracture across the column were recorded, along with a measure of the fractured weak layer depth and its date ID if applicable.

Frequently, damping snow (the compressed snow between the shovel and the fractured weak layer) was measured and recorded, as was the length of any propagation that didn`t reach the column end, and the observed `fracture character` (van Herwijnen and Jamieson, 2004a) in the layer. Although fractures initiated in numerous storm layers without propagation in the ECT, these results were infrequently recorded, as the objective was to compare ECT results to PST results which focused on specifically selected (and primarily persistent) weak layers.

Figure 3.5: The operator isolates the ECT column from the surrounding snowpack via cord-cut to a desirable depth (a) then loads one end of the column with 10 taps from the wrist, 10 from the elbow, and 10 from the shoulder (b), noting all fractures that initiate and propagate within the column.

The recording standard used to present the results was as proposed by Simenhois and Birkeland (2007, 2009) and is as follows, where *n* is the number of taps:

- **ECTPV** – fracture propagates across the entire column during isolation;
- **ECTP *n*** – fracture initiates on the n^{th} tap and propagates across the entire column on the n^{th} or n^{th}+1 tap;
- **ECTN *n*** – fracture initiates on the n^{th} tap, but *does not* propagate across the entire column on the n^{th} or n^{th}+1 tap;
- **ECTX** – a fracture does not initiate within the 30 taps (given a score of 35).

Propagation is said to be *likely* (*propL*) on surrounding slopes in the case of ECTPV and ECTP *n*, provided *n* is within the 30 standard taps. Otherwise, propagation on surrounding slopes is said to be *unlikely* (*propUL)*.

Errors in performing the ECT were minimal. Column lengths within approximately 10% of the recommended 90 cm were accepted. To ensure consistent results when loading the column (tapping), the same operator performed all ECTs on a given field day and ensured the shovel blade did not overlap into the undisturbed snowpack surrounding the isolated column.

3.5 Validation Sites

The method of skiing a small slope to test its stability (ski-cutting) is shown in Figure 3.6 and has been used in numerous studies to validate a variety of snowpack stability tests (e.g. Jamieson and Johnston, 1993a; Simenhois and Birkeland, 2007) and profile interpretation (e.g. Schweizer and Jamieson, 2003). Typically, a slope is rated unstable if a skier-triggered avalanche is released (or if the site previously avalanched naturally or whumpfed). A slope is rated stable if no avalanche released, often after multiple attempts to trigger it. This 'stable' rating is somewhat subjective because there is no evidence as to whether a fracture in the primary weak layer was initiated during the ski-cut and therefore uncertainty remains as to whether that layer may have propagated *if initiated*, perhaps from a shallower spot elsewhere on the slope. Simenhois and Birkeland (2007) acknowledged this potential source of error when they noted that ski-cuts may not have initiated fracture in the layer that then propagated in the ECT elsewhere on the slope, generating a significant number (16%) of false-unstable predictions on stable slopes. For this study, a similar method for rating unstable sites was used: any sign of instability *and* propagation observed on the slope/flats or in the immediate surroundings (adjacent slopes typically within 100 m of the selected site) indicated an unstable site (Figure 3.7), including natural avalanches, remotely triggered avalanches, whumpfs, or skier-triggered avalanches – provided obvious propagation was evident and a slab released or whumpfed.

Figure 3.6: A controlled method of skier-testing a small slope, called ski-cutting, is performed by skiing across the upper portion of the slope and down-weighting near the point of greatest convexity. Ski-cutting is an attempt to initiate a fracture within a buried weak layer which may propagate to cause a small avalanche.

However, for rating stable slopes the ski-tracks were dug up after each ski-cut that *did not* release a slab to determine whether the ski-cut had *initiated* fracture in the weak layer and that the fracture just had not propagated (Gauthier, 2007, pp. 89; Gauthier and Jamieson, 2008a) (Figure 3.8). If the weak layer appeared undisturbed after the ski-cut, it was not concluded that the site was stable. If the weak layer *had* been crushed or obviously disturbed, it was concluded that the weak layer and slab combination at that site had been given the *chance* to propagate, but appeared to lack propagation propensity. The site was therefore rated stable. This method allowed direct comparison of PST results with observed propagation propensity on the slope scale at *Validation Sites*.

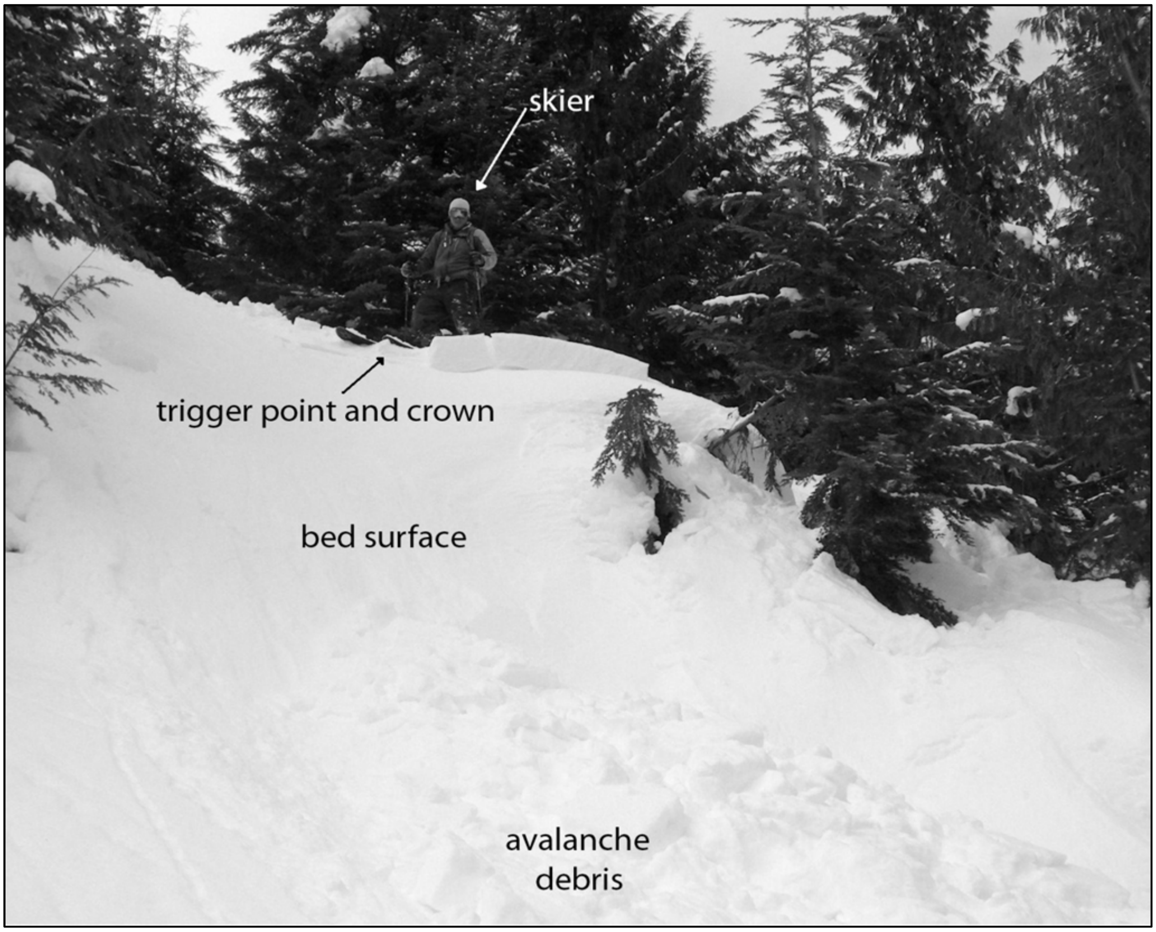

Figure 3.7: A skier-cut avalanche above the Fidelity access road used as a validation site in February of 2009. The failure layer was approximately 40 cm deep and was triggered by the skier at the top of the slope. The same layer was tested nearby with the standard PST and short-scaled PST.

During the 2008 and 2009 winters, standard PST tests were performed at 53 validation sites, supplementing the 22 sites with standard PSTs from 2006-07 previously available to validate the PST (Gauthier, 2007; Gauthier and Jamieson, 2008a). At some validation sites two PWLs were fractured during a ski-cut or in a natural avalanche of which neither, one, or both showed high propagation propensity. In this case, different stability ratings could be assigned to each layer, called *validation site-layers*, and later compared to test results on that specific layer.

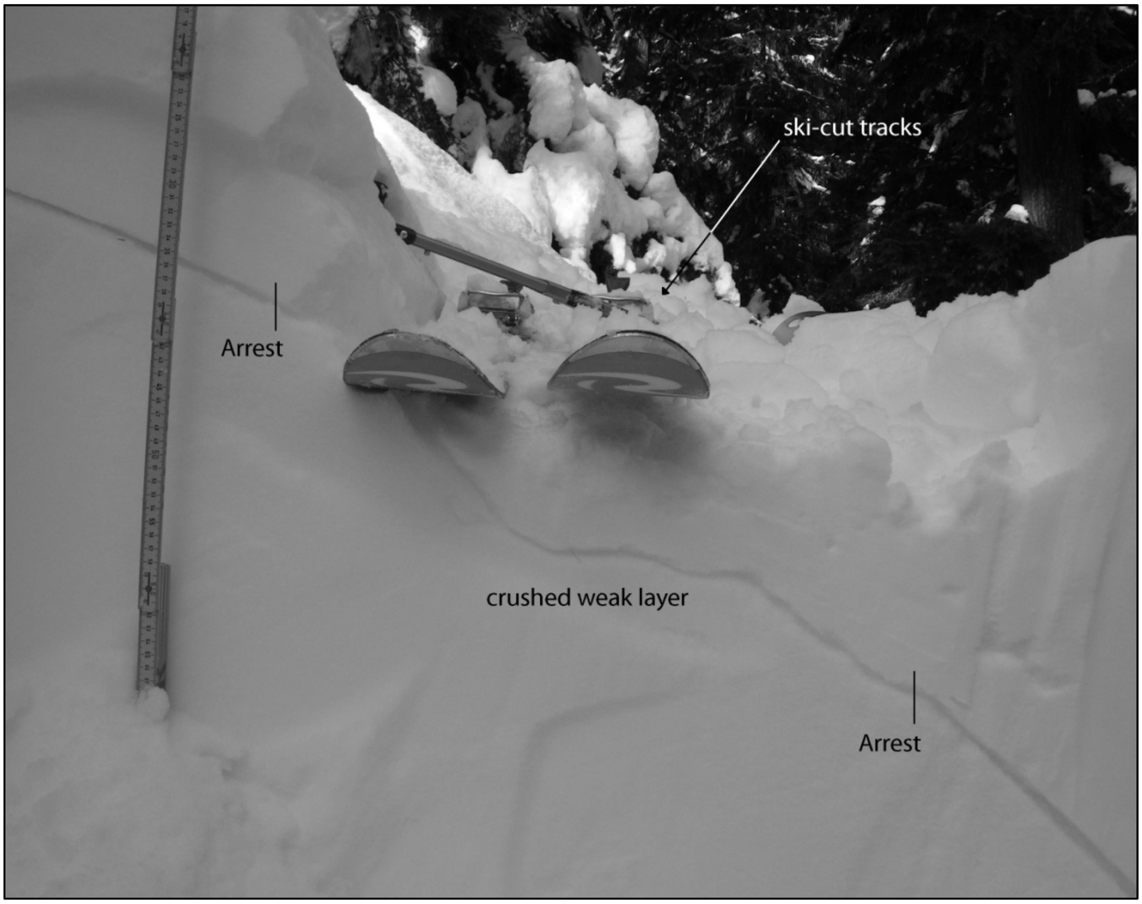

Figure 3.8: Exposed ski tracks after a ski-cut attempt failed to release an avalanche or whumpf. Note that below the skis the weak layer has been crushed and deformed, while both up-slope and down-slope from the skis the weak layer remains intact. Since a fracture was initiated but arrested without propagation, this site was deemed stable for this study.

3.6 Short-Scaled PSTs

Short-scaled PSTs were performed for this study in the winter of 2009 and were identical in test and recording method to the standard PST (section 3.3) with the exception that the column length was scaled to match weak layer depths *less than* 1 m as illustrated in Figure 3.9. Since the standard method already employs a scaled column length to match weak layer depth beyond 1 m, the short-scaled columns only apply to layers less than 1 m deep.

Gauthier and others (2008) originally suggested that scaling the column length with weak layer depth under a metre, and not just above, might reduce the number of false-stable results. This was in light of their observed higher number of false-stable results in longer test columns (e.g. 1.5 m) and soft, thin slabs during early experimentation with the PST. Although the intention of scaling the PST below 1 m was to potentially reduce the number of false-stable results observed in shallow, soft slabs (Gauthier et al., 2008), a full range of column lengths from 10 – 100 cm was tested over the season in case results led to a permanent change in the standard PST test method.

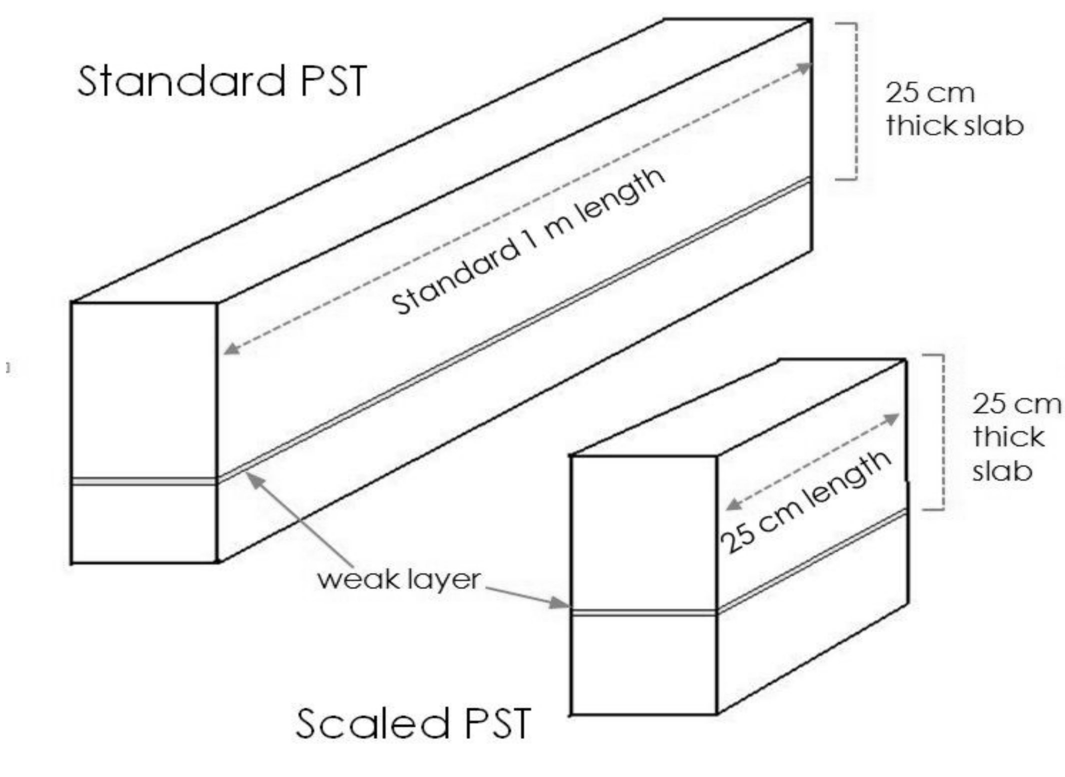

Figure 3.9: The standard 1 m long PST compared to the short-scaled PST, using a column length scaled to match weak layer depths below 1 m. For layers deeper than 1 m, the standard PST is already scaled in length to match.

3.7 Validation of Deep-Slab PSTs

To assess the validity of the PST in predicting propagation propensity in deeply buried weak layers where initiating fractures by skier-testing was either unsafe or impractical, a method other than using validation sites had to be assumed. Independent of ASARC test results, the lead forecasters in the Avalanche Control Section (ACS) of Glacier National Park were asked to rate the propagation propensity for each PWL buried in the snowpack on a daily basis, to which PST results could be compared. These expert ratings were based on field observations and avalanche control results obtained by the ACS and involved rating each PWL as *unlikely, equally likely,* or *likely* to propagate fractures far enough to release an avalanche if initiated. In addition, the expected elevation band(s) of below tree line (BTL), at tree line (TL), or alpine (ALP) for which the rating applied were noted, along with an estimate of the extent of propagation expected in the avalanche start zone (i.e less than the start zone size, < SZ; the approximate start zone size, SZ; or greater than the start zone size, > SZ).

For deeply buried weak layers that were unsafe or impractical to trigger in the field, PST results could potentially be compared to the expert ratings of propagation propensity for that layer on that day. Although the expert ratings applied to the expected average condition of the specific slab and weak layer combinations through the entire park on that day rather than at that day's test site, the chosen sites – especially for deeply buried layers – were considered representative of the average snowpack conditions in the park, particularly those sites at Mt. Fidelity. If initial comparisons showed a strong correlation between expert ratings of propagation propensity and PST results in the same elevation bands, further correlation could be sought between the expected extent of propagation and propagation lengths observed in the PST.

3.8 Regional-scale Validation of the PST

Daily avalanche activity recorded over the 2008 and 2009 winters throughout the Columbia Mountains was used in attempt to validate the PST on the regional scale. The source for these data was the CAA-operated Industry Information Exchange system known as InfoEx, used under a conditional data-licence agreement drafted by the CAA.

The InfoEx is a database of technical snow, weather and avalanche occurrence information updated daily through submissions from subscribing operations in different regions of BC and Alberta. The same operations then use the InfoEx to synthesize regional avalanche activity information to help forecast local avalanche hazard. Subscribers are primarily ski resorts, parks, highway operations and ski guiding operations. The recorded avalanche observations include avalanche size, type of trigger, weak layer depth and grain type, estimated slab dimensions, and start zone elevation. Size is based on the Canadian size classification of 1 -5 (CAA, 2007) as given in Table 3.1, and includes half-sizes (e.g. 3.5) if the avalanche appeared to be between two size classifications. Based on avalanche activity reported in the InfoEx on specific PWLs within the Columbia Mountains, the PST results obtained on those same layers in Glacier National Park and around Blue River could be assessed in terms of evaluating propagation propensity at the regional scale.

Avalanche observations from 39 operations in the Columbia Mountains were included, all of which reported avalanche activity on the same PWLs (grain type and approximate burial date) as were being tested around Blue River and in Glacier National Park. Although this encompasses an estimated area of over 30,000 km^2, justification for this spatial scale decision is based on the work of Hägeli and McClung (2003, 2007) and is described further in Section 3.8.1.

The relevant information extracted from the InfoEx database included the avalanche occurrence dates, sizes, and weak layers on which the avalanche released. Other criteria used to filter the avalanche observations are described in Section 3.8.2.

Table 3.1: Avalanche Size Classification (CAA, 2007). Half-sizes are often reported. Note that the typical mass increases by a factor of 10 with each size increment.

Size Class	Destructive Potential Definition	Typ. Mass
1	Relatively harmless to people.	< 10 t
2	Could bury, injure, or kill a person	10^2 t
3	Could bury or destroy a car, damage a truck, destroy a wood-frame house or break a few trees.	10^3 t
4	Could destroy a railway car, large truck, several buildings or a forest area of approximately 4 hectares.	10^4 t
5	Largest snow avalanche known. Could destroy a village or a forest area of approximately 40 hectares.	10^5 t

Only PSTs from assumed regionally representative study sites at Mt. Fidelity and Mt. Abbott in Glacier National Park, and at Mt. Saint Anne near Blue River, were evaluated next to the regional avalanche activity to determine if propagation trends existed. These sites have been used historically by the respective operations and by researchers since their weather and snowpack are representative of conditions at tree line in the surrounding area (Section 3.1) – although not necessarily areas outside the operation, let alone the entire Columbia Mountains. Although snowpack tests have shown variability within individual slopes (e.g. Campbell and Jamieson, 2007; Hendrikx et al., 2009), and cannot be expected to definitively indicate stability (or propagation propensity) over a large scale (e.g. Hägeli and McClung, 2004), an aggregate of tests at representative sites *may* correspond with trends in propagation propensity seen throughout the region. However, because of these scale issues, it is expected that any correlations between PST results and regional avalanche activity will not be strong, and cannot be as good as validated predictions for nearby slopes.

3.8.1 'Avalanche climate' of the Columbia Mountains

Although the Columbia Mountains host more than 50 reporting operations throughout a huge part of interior BC (Figure 3.10), Hägeli and McClung (2003, 2007) showed that most significant PWLs are the result of widespread regional weather events, cover large areas, and produce avalanche activity across the entire mountain range. They observed that avalanche activity could be more pronounced in specific areas, but found no statistical evidence indicating that the number of PWLs or associated avalanche activity were functions of geographical location within the Columbia Mountains.

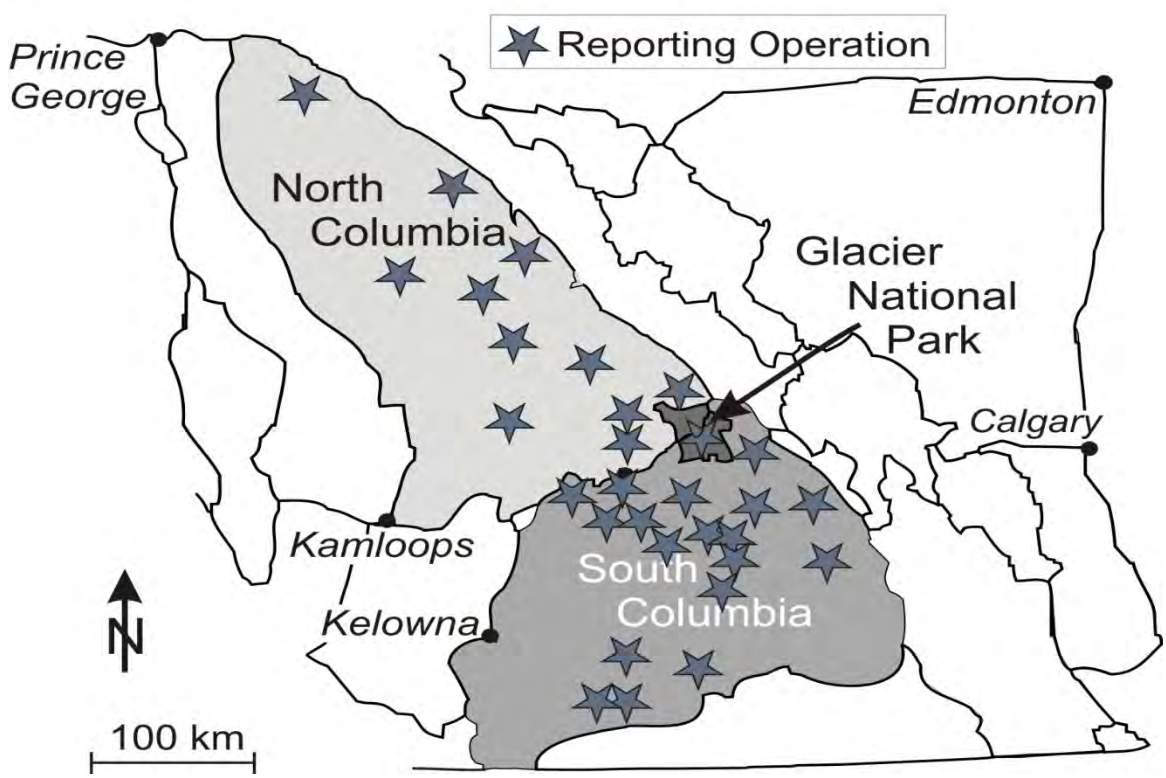

Figure 3.10: Map of south-western Canada showing the CAC forecast regions of the North and South Columbias, including Glacier National Park, and showing the approximate location of the 39 operations that reported avalanche activity on at least one of the PWLs tested with the PST for this study (after Jamieson et al., 2008). Although the Columbia forecast regions extend west to Kamloops and Kelowna, the majority of operations are in the eastern half of the regions.

Both the 2008 and 2009 winters exhibited patterns that are typical of the Columbia Mountains according to Hägeli and McClung (2003), including an early-season crust-facet layer and three successive surface hoar layers spread throughout the season. In fact, the same authors found that most operations report at least one early-season crust-facet layer and one to three persistent surface hoar layers each year with no significant north-south or east-west variation in numbers. Additionally, they found that early-season crust-facet layers produced intermittent avalanche activity throughout the season while each surface hoar layer typically produced avalanches over three to four weeks. Similar trends were observed in all eight PWLs throughout 2008 and 2009.

3.8.2 Avalanche activity index (AAI)

Regional avalanche activity from the Columbia Mountains was compiled according to the following criteria and then used in the calculation of a daily avalanche activity index (AAI; equation 3.1):

- Only dry slabs were included. Loose and wet avalanches were excluded;
- Avalanches without occurrence dates, estimated size, and/or an identified PWL were excluded;
- Only avalanches of size 1.5 and larger were included (e.g. Jamieson and Johnston, 1993b). Size 1 avalanches occur frequently and are relatively harmless, are generally recorded with less consistency in the InfoEx, and carry little weight in the calculated avalanche activity index;
- All types of trigger were included (natural and artificial);
- All aspects and elevations were included.

Since burial date IDs vary across the region by a few days depending on the arrival of storm snow or operational naming procedure, careful selection of dates on either side of the local burial date helped compile avalanche activity on the eight major PWLs of the past two seasons. Weak layer grain-type was checked to ensure the dated layers were in fact the same PWL being tested with the PST. Only 37% of reported avalanche activity in the InfoEx over the 2008 and 2009 winters had identified weak layer dates and 20% of reported avalanches were on the eight PWLs commonly tested. This is comparable to the observations of Hägeli and McClung (2007).

Based on the number and sizes of avalanches reported in the InfoEx on a given day, a daily avalanche activity index (AAI) could be calculated for the entire Columbia Mountains region. The AAI equation was adopted from Jamieson and others (1998) and is as follows:

$$AAI = \sum N_i \, 10^{i-1} \tag{3.1}$$

Where *i* is the avalanche size class, including half-sizes, and N_i is the number of avalanches of size *i* on a given day. The AAI gives a ten-fold increase in weighting to each full size class, analogous to the typical mass increase between size classifications. The number of avalanches is perhaps influenced more by triggering ease than propagation propensity, although propagation is obviously required for a slab avalanche of any size. The size of an avalanche is a more direct indication of propagation propensity, and thus the AAI gives greater weight to large avalanches that show extensive propagation across terrain. For example, 10 size-two avalanches would have an AAI of 100, equivalent to one size-three avalanche. A size-four avalanche would have an AAI of 1000.

Often, avalanche observations are made in poor weather conditions and from far distances. This leads to missing information and potential errors in identifying the failure layer and, less frequently, the avalanche size. Additionally, many observations are made up to a few days after an avalanche occurred, leading to estimates of occurrence dates. This is the result of operations (except ski-hills) typically observing only a fraction of their terrain on a given day (e.g. Jamieson, 1995, pp. 61; Hägeli and McClung, 2007), which depends on weather conditions and the number of field investigations, highway patrols, or groups skiing within the terrain. Although these errors in estimating and missing information have a small affect on the calculated AAI on a given day, the AAI is expected to reflect the trends in avalanche activity in the Columbia Mountains over multiple days and throughout the season.

Occasionally during large avalanche cycles, operations will report ‘several’ or ‘numerous’ avalanches of the same size on the same PWL within the same day. In order to calculate the AAI, ‘several’ and ‘numerous’ were converted to the numerical values of 5 and 10, which is similar to conversions by Hägeli and McClung (2003) and Jamieson and others (1998). The term ‘isolated’ is also commonly entered as the number of avalanches which, according to the accompanying comments in all cases, refers to a single avalanche.

4. RESULTS

The results of two winters of fieldwork with the PST and one with the ECT are presented in this chapter. The objectives for this work were outlined in Section 1.10 and the field methods for gathering data were described in Chapter 3. A summary of the data collected for this study is presented in Section 4.1. Section 4.2 updates the validation study on the PST by combining data collected as part of this study with an earlier dataset (see Section 3.5). Section 4.3 assesses the results of scaling the PST column length below 1 m. Comparative results of the PST and ECT from within the same snowpits are presented in Section 4.4, indentifying trends, advantages, and limitations of each test beyond the validation results. Preliminary validation of the PST for deep slabs using expert ratings of propagation propensity is presented in Section 4.5, while Section 4.6 presents the results of regional validation comparing PST results from study sites to avalanche activity throughout the Columbia Mountains.

4.1 Dataset

The PST dataset analysed for this study was collected throughout the winters of 2008 and 2009. ECT data are entirely from 2008 and short-scaled PST data are from 2009. The validation dataset draws from a previous season's data (2007) before the author was present, but only columns meeting the standard test method (Section 3.3) are used. Two seasons of expert ratings (2008, 2009) exist for validating the PST in deep slabs within Glacier National Park (Section 3.7). Additionally, PST data and regional avalanche activity were tracked over the past two seasons to compare the growth and decay of propagation propensity in specific persistent weak layers within the test itself and throughout the Columbia Mountains.

In total, 783 standard PSTs were performed in the winters of 2008 (412) and 2009 (371) in 258 snowpits on 193 field days. 246 ECT results were recorded in 183 individual tests performed in 91 snowpits in 2008. In 2009, 231 short-scaled PSTs ranging in slab thickness from 10 cm to 100 cm were performed beside standard PSTs. Table 4.1 summarizes the complete dataset, of which various subsets were used in the analysis. For instance, 365 of the 412 PSTs and 242 of the 246 ECTs performed in 2008 had the tested layer depth recorded as required for depth-related analysis, and only 169 PSTs have been performed next to validation sites.

Table 4.1: Complete dataset of PSTs, ECTs, and short-scaled PSTs collected in 2008 and 2009 for the current study.

<table>
<tr><th colspan="2">Individual Test Results</th><th>Rogers Pass</th><th>Blue River</th><th>Kicking Horse</th><th>Chatter Creek</th><th>TOTAL</th></tr>
<tr><td rowspan="2">2008</td><td>ECT*</td><td>101</td><td>127</td><td>12</td><td>6</td><td>246</td></tr>
<tr><td>PST</td><td>215</td><td>178</td><td>13</td><td>6</td><td rowspan="2">783</td></tr>
<tr><td rowspan="2">2009</td><td>PST</td><td>203</td><td>154</td><td>14</td><td>0</td></tr>
<tr><td>short-PST</td><td>133</td><td>86</td><td>12</td><td>0</td><td>231</td></tr>
</table>

*Includes multiple results in single test columns

4.2 Slope-scale Validation of the Propagation Saw Test

Performing the PST at sites of confirmed propagation or confirmed initiation *without* propagation (validation sites, Section 3.5) enabled the predictive accuracy of the PST to be evaluated at the slope scale. This is essential to establishing the credibility of any field test, and up to this point validation results for the PST have been limited (see Table 2.2). Furthermore, Gauthier's (2007, pp. 91) unique method of examining ski-cut tracks for fracture initiation without propagation explicitly allowed propagation to be verified. The results presented here draw from the 2008 and 2009 datasets and include the validation data gathered using the standard test method (Section 3.3) in 2006-07 and presented in part by Gauthier (2007), Gauthier and others (2008) and Gauthier and Jamieson (2008a, 2008b).

4.2.1 Data

During the winters of 2007 to 2009, standard PSTs were performed at 75 validation sites. Since some of these sites had multiple layers that fractured during the same event (e.g. ski-cut; Figure 4.1), a total of *84 instances* of fractured layers, called *validation site-layers*, were tested at the 75 different sites. The full list of validation site-layers and the associated PSTs performed on them can be found in Appendix A.

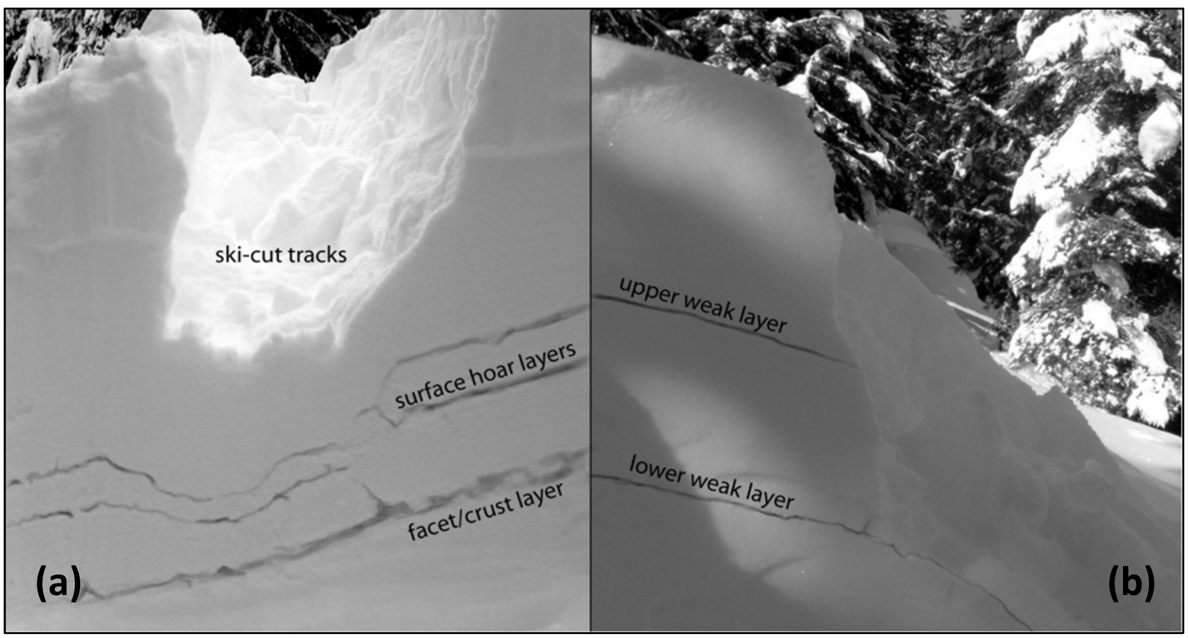

Figure 4.1: At some validation sites, multiple layers fractured in the same ski-cut or avalanche event and could each be tested with the PST. In (a), three layers are present in the snowpack, in which fractures in the upper two were initiated by the ski-cut, but neither propagated. The lower layer was undisturbed and could not be tested for the validation study. In (b), fractures in two layers were initiated and the only the lower one showed obvious propagation. The whole slab above the lower layer slid but a fracture did not propagate along the upper layer.

Table 4.2 summarizes the validation dataset gathered between 2007 and 2009, including the number of validation sites, the number of validation site-layers by incident type (avalanche/whumpf, or fracture initiation without propagation), and the number of standard PSTs performed by incident type. Field workers made greater efforts in 2009 to investigate the ski tracks of unreleased ski-cuts, evident from the increased number of tests on layers of confirmed initiation without propagation. In total, 119 tests (70%) were performed on 59 site-layers showing confirmed propagation, and a further 50 tests (30%) were performed on 25 site-layers showing confirmed initiation without propagation (Figure 4.2). This resulted in 169 validated standard PSTs.

Table 4.2: The PST-validation dataset gathered between 2007 and 2009. Only test columns meeting the standard length described in Section 3.3 are included and only at sites of confirmed initiation with or without propagation.

SEASON:	**2007**	**2008**	**2009**	**2007-2009**
Total Validation Sites	22	25	28	**75**
Validation site-layers by incident type:				
Confirmed propagation (avalanche/whumpf)	17	25	17	59
Confirmed initiation without propagation	5	3	17	25
Total validation site-layers:	22	28	34	**84**
Number of PSTs performed by incident type:				
Confirmed propagation (avalanche/whumpf)	29	52	38	119
Confirmed initiation without propagation	9	5	36	50
Total validated PSTs	38	57	74	**169**

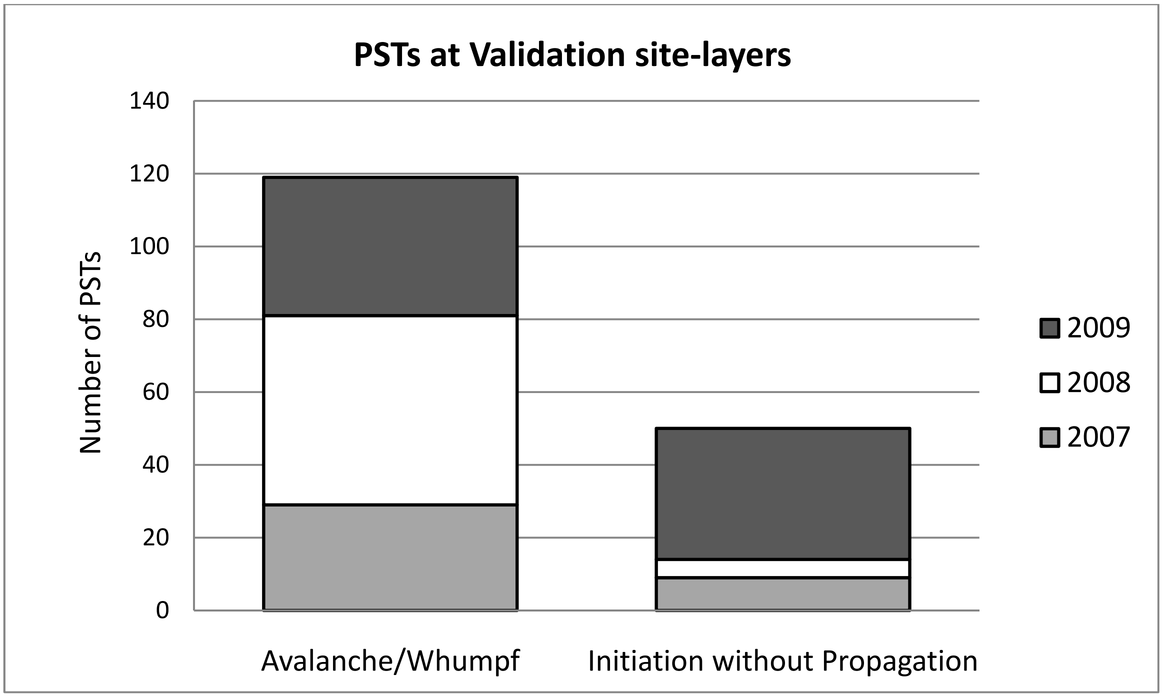

Figure 4.2: Total number of PST results gathered over three winters, separated by incident type (confirmed propagation or confirmed initiation without propagation).

4.2.2 Slope-scale validation results

PST results that indicated propagation was likely (*propL)* next to an avalanche or whumpf were classified as correct predictions, as were test results that indicated propagation was unlikely (*propUL*) next to a site of confirmed initiation without propagation. False-stable PST results were those that predicted propagation was unlikely next to an avalanche or whumpf. False-unstable results predicted propagation was likely next to a site of confirmed initiation without propagation. The terms 'stable' and 'unstable' are used here as descriptors of 'low or no' and 'high' propagation propensity, respectively (described in Section 1.7.2), as opposed to their more traditional use of describing stability in terms of fracture initiation and propagation together. The validation results by year are summarized in Table 4.3 and the cumulative results for the three seasons are tabulated in a contingency table commonly used to objectively evaluate the accuracy of a predictive tool or forecast (e.g. Doswell and Flueck, 1989).

Table 4.3: Validation results from the winters of 2007 to 2009, indicating the number of tests at validation sites that gave correct predictions of propagation potential and the number that gave incorrect predictions. The cumulative results are re-presented in the form of a contingency table commonly used to evaluate predictive tools.

SEASON:		2007	2008	2009	2007-2009
Prediction Type	Correct Stable	9	4	32	45
	Correct Unstable	22	31	28	81
	False-Stable	7	21	10	38
	False-Unstable	0	1	4	5
Total:		38	57	74	169

PROPAGATION:		OBSERVED YES	OBSERVED NO
PREDICTED	YES	81	5
	NO	38	45

2007-2009 Validation Contingency Table (n = 169)

One way to look at these results is to evaluate the overall frequency of each prediction type as a percentage of all validated PSTs. The prediction-type frequencies from individual seasons are compared in Table 4.4, and the overall prediction accuracy for the current validation dataset is shown in Figure 4.3.

Table 4.4: Predictive accuracy of the PST validation dataset presented by year. Results are mostly comparable over the three years, although 2008 had a substantially higher number of false-stable results, discussed in Section 4.2.4

SEASON:	2007	2008	2009	2007-2009
Correct Predictions	31 (82%)	35 (61%)	60 (81%)	126 (75%)
False-Stable	7 (18%)	21 (37%)	10 (14%)	38 (22%)
False-Unstable	0	1 (2%)	4 (5%)	5 (3%)
Total	38	57	74	169

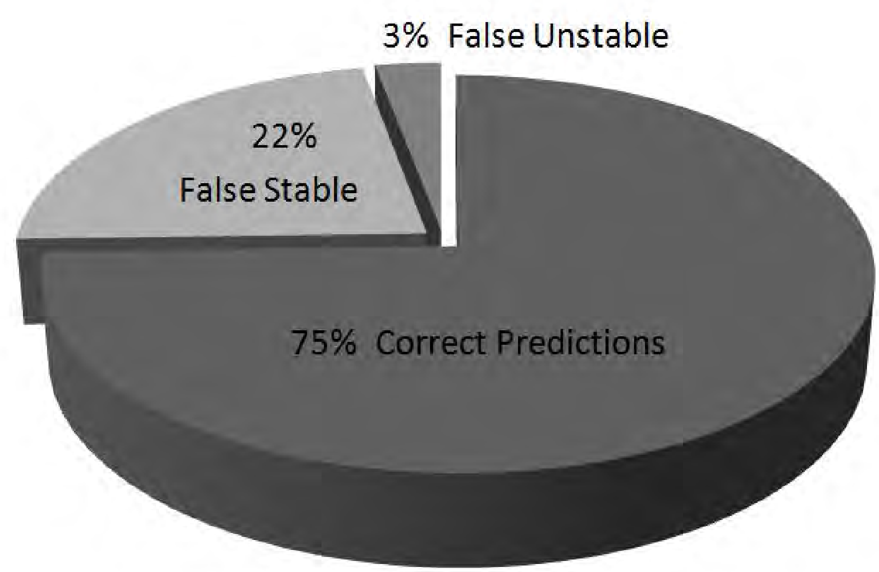

Figure 4.3: Overall predictive accuracy of the full validation dataset as percentages of all validated PSTs.

Validation results show comparable prediction-type frequencies year to year, with particularly similar frequencies in 2007 and 2009 where the overall accuracy (frequency of correct predictions) was over 80% and false predictions were relatively low. In 2008, a higher number of false-stable predictions were observed which brings the overall accuracy of the dataset to 75% and the frequency of false-stables to 22%. False-unstable predictions were consistently low, comprising only 3% of the total dataset.

4.2.3 Contingency tables and measures of predictive skill

The contingency table (Table 4.3) facilitates the calculation of numerous statistics that provide different perspectives on the results of a dichotomous forecast. Table 4.5 gives the names and equations of these statistics (after Doswell et al., 1990) based on the general 2 x 2 contingency table also shown. The validated PST results and their calculated ratios are given in parentheses.

Table 4.5: a) The general 2 x 2 contingency table and b) all the common ratios calculated from it (after Doswell et al., 1990). These statistics can be used to gain different perspectives on the value of a predictive tool such as the PST, for which the validation data and calculated ratios are given in parentheses.

Table 4.5 a) The 2x2 contingency table

		OBSERVED	
		YES	NO
PREDICTED	YES	*a* (81)	*d* (5)
	NO	*c* (38)	*b* (45)

POD: Probability of Detection
FAR: False Alarm Ratio
FOM: Frequency of Misses
PON: Probability of a Null event
FOH: Frequency of Hits
POFD: Probability of False Detection
DFR: Detection of Failure Ratio
FOCN: Frequency of Correct Null predictions

Table 4.5 b) Statistic calculations

POD = $\frac{a}{a+c}$ (= 0.68)	FOH = $\frac{a}{a+d}$ (= 0.94)
FOM = $\frac{c}{a+c}$ (= 0.32)	DFR = $\frac{c}{c+b}$ (= 0.46)
FAR = $\frac{d}{a+d}$ (= 0.06)	POFD = $\frac{d}{d+b}$ (= 0.10)
PON = $\frac{b}{d+b}$ (= 0.90)	FOCN = $\frac{b}{c+b}$ (= 0.54)

These statistics give insight into the predictive success of the current PST dataset based on relative comparisons between various predictions and observations. For example; when the PST predicted propagation, the same layer propagated a fracture when initiated on an adjacent slope 94% of the time (FOH). In only 10% of cases the PST falsely predicted propagation when conditions were stable (POFD). Conversely, next to an unstable slope, the PST falsely predicted low or no propagation propensity 32% of the time (FOM). Perhaps the most important statistic is the probability of detection (POD) which indicates the percent of unstable slopes that are detected. For the current PST validation data, the POD was 68%.

The True Skill Statistic (TSS) method (e.g. Doswell et al., 1990) of evaluating forecasts (predictions) provides a relative measure of prediction skill based on two of the simple statistics calculated from the contingency table. The term 'skill' is used instead of accuracy when comparing the relative success of the data to a baseline or standard forecast; in this case, a hypothetical perfect forecast where there are no false predictions (*c* = *d* = 0).

The TSS equation is:

$$\mathrm{TSS} = \mathrm{POD} - \mathrm{POFD} = \frac{(\mathrm{ab} - \mathrm{cd})}{(\mathrm{a} + \mathrm{c})(\mathrm{d} + \mathrm{b})} \tag{4.1}$$

where the hypothetical perfect forecast would have a TSS equal to 1.0. The TSS equation calculates a skill value between -1 and 1 by subtracting the probability of false detection from the probability of detection (for which the perfect forecast would have probabilities of 0 and 1, respectively). A negative TSS value indicates a worse result than a randomly generated set of predictions. The TSS for the current PST validation dataset is:

$$\mathrm{TSS} = (\mathrm{POD} - \mathrm{POFD}) = (0.68 - 0.10) = 0.58$$

Statistics calculated from contingency tables, specifically the TSS, have been employed in the comparison of snowpack stability tests by Gauthier (2007), Gauthier and Jamieson (2008b), Simenhois and Birkeland (2009) and Moner and others (2008). An alternative but closely related measure of skill, employed by Winkler and Schweizer (2009) and Schweizer and Jamieson (2009), involves combining the sensitivity (POD) and specificity (PON) in an unweighted average accuracy (UAA) where a perfect forecast again equals 1.0:

$$\mathrm{UAA} = \frac{\mathrm{POD} + \mathrm{PON}}{2} = \frac{\mathrm{sensitivity} + \mathrm{specificity}}{2} \tag{4.2}$$

which for the current PST data equates to:

$$\text{UAA} = \frac{\text{POD} + \text{PON}}{2} = \frac{0.68 + 0.90}{2} = 0.79$$

It can be argued that the UAA (or TSS) is more appropriate than overall accuracy when the dataset has an unbalanced number of tested stable and unstable slopes (e.g. Doswell et al., 1990; Winkler and Schweizer, 2009), as it does in this study (n_{stable} = 50, $n_{unstable}$ = 119). A test with a high sensitivity accurately predicts unstable events (high propagation propensity) in most cases, which is ideal. A high specificity (low false-unstable rate) is also preferred since false alarms can lead to a loss of credibility. It is important to note that neither the TSS nor UAA weigh false-stable results more heavily than false-unstable results, even though the former are generally considered more dangerous and therefore less desirable.

4.2.4 False predictions of the PST

False-stable results are more dangerous and consequently less desirable than false-unstable results for obvious reasons: they give a false prediction of stability – or low propagation propensity in the case of the PST – which can lead practitioners and recreationists to underestimate the existing hazard and make the potentially high-consequence decision to enter avalanche terrain. That being said, false-unstable results are also undesirable since they reduce the test's credibility if too many false alarms are forecast. The PST has shown a low number of false-unstable results over the three seasons of validation, but has shown a high number of false-stable results. In early validation studies (Gauthier and Jamieson, 2007b), false-stable results were commonly associated with shallow, soft slabs. In 2009, four such false-stable results occurred on three site-layers, all less than 25 cm deep, with soft overlying slabs (F, F-).

Additionally in 2009, five false-stable results were on the flanks or above the crown of two avalanches that had occurred the day prior to testing. In 2008, 14 of 21 false-stable results occurred on the flanks or above the crown of day-old avalanches. This suggests that the tested areas may have actually been stable, or that changes in the surrounding snowpack after an avalanche may have occurred quickly and, in some cases, tests performed at the site the following day no longer showed the high propagation propensity the slope previously had. These false-stable cases are omitted in a re-calculation of predictive accuracy and skill in Section 4.2.5, and discussed further in Section 4.2.8.

The small number of false-unstable results have no distinguishing characteristics other than that they were all in surface hoar layers, and were all accompanied by at least one accurate prediction (see Section 4.2.6). In fact, the five false-unstable results at three different sites had weak layer depths that ranged from 30 to 65 cm deep and cut lengths between 22% and 42% of the column – well below the 50%-cut threshold. The small number of false-unstable results is probably too few to establish any kind of significant relationship with potential contributing factors such as grain type or layer hardness.

4.2.5 Adjusted validation next to avalanches and whumpfs

Removing the 12 validation site-layers that were tested the day after they avalanched reduced the number of false-stable results by 19, and also reduced the number of correct unstable predictions by 8, to give the results shown in Table 4.6. Doing so increased the overall accuracy from 75% to 83%, the TSS from 0.58 to 0.69, and the UAA from 0.79 to 0.85. False-stable predictions were reduced from 22% to 14% overall.

Table 4.6: Adjusted validation results for the PST in which 12 site-layers tested the day after they avalanched were removed. Both correctly and incorrectly predicted site-layers were removed, reducing the number of correct predictions by 8 and false-stable results by 19 (n =140).

PROPAGATION:		OBSERVED	
		YES	NO
PREDICTED	YES	73	5
	NO	19	45

Prediction Strength
Accuracy: 83 %
TSS: 0.69
UAA: 0.85

4.2.6 Reproducibility of test results at validation sites

Eleven of seventy-five sites (15%) in the validation dataset had both correct and incorrect PST predictions on the same layer side by side in the same snowpit. All three stable sites with one or more false-unstable PST result also had at least one correct test beside it. The remaining eight sites were unstable with both correct and false-stable predictions, which comprises more than one third of all the sites in the validation dataset where false-stable predictions were encountered (8 of 23 sites). Using the 2007 validated dataset and including non-standard-length columns, Gauthier (2007, pp. 132) experienced both correct and incorrect predictions at 13 of 23 validated sites (56%). This exemplifies the improved reproducibility of results when only the standard column length is used.

4.2.7 Comparison with other standard snowpack tests

Table 4.7 compares the overall accuracy, TSS and UAA for the current PST validation dataset with that of other common snowpack tests. TSS and UAA values for other tests were calculated from a sample-weighted average of multiple datasets summarized by Schweizer and Jamieson (2009), as was the overall accuracy of the Threshold Sum. Other overall accuracies are from single sources (Gauthier and Jamieson, 2008b; Simenhois and Birkeland, 2009). The reader is directed to those works and that of Schweizer and Jamieson (2009) for descriptions and sources of the datasets used to calculate values for the other tests. It is important to note that the values reported for the other tests measure the success of predicting fracture initiation and propagation together (Gauthier and Jamieson, 2008b), since they are typically compared to slopes that were skier tested and either avalanched or not. In the case of the PST, where initiation was confirmed in all validation cases, only the success of predicting propagation was measured. Table 4.7 includes the adjusted PST validation results with day-old avalanches omitted, which are given in parentheses since the justification for excluding them is unproven and is unique to this study.

Table 4.7: Comparison of the predictive strength of the PST to other common snowpack tests and indicators in three related categories: overall accuracy, True Skill Statistic (TSS), and unweighted average accuracy (UAA).

Snowpack Test:	***n***	**Overall Accuracy**	**TSS**	**UAA**
PST	**169**	**75 %**	**0.58**	**0.79**
(PST w/o day-old avalanches)	(140)	(83%)	(0.69)	(0.85)
ECT	635[3], 582[1]	93 % [3]	0.75 [1]	0.88 [1]
Threshold Sum	599[1]	67 % [1]	0.30 [1]	0.65 [1]
CT score	285[1]	~	0.35 [1]	0.68 [1]
CT score + fracture character	58[2], 179[1]	79 % [2]	0.46 [1]	0.73 [1]
RB score	23[2], 828[1]	69 % [2]	0.48 [1]	0.71 [1]
RB score + release type	23[2], 364[1]	70 % [2]	0.53 [1]	0.76 [1]

Sources: [1] sample-weighted average from Schweizer and Jamieson (2009); [2] Gauthier and Jamieson (2008b); [3] Simenhois and Birkeland (2009) combined Colorado-New Zealand and SnowPilot datasets.

As evident in Table 4.7, the predictive strength of the PST is comparable to other standard snowpack assessment methods. The overall accuracy is slightly higher than the rutschblock and only four percentage points below that of the compression test, and the PST has a higher TSS and UAA than all other methods except for the ECT, although statistical significance was not tested. This is strong evidence for the validity of the PST, since tests like the RB and CT have been in use by practitioners and recreationists for decades.

Gauthier and Jamieson (2008b) recalculated the TSS values for the CT, RB, and Threshold Sum for only sites where initiation was confirmed (therefore only propagation is predicted), and showed that all three had more false-unstable results (greater POFD) and therefore lower TSS values at these sites. In fact, compared to the TSS of 0.58 for the PST determined here, the propagation predictions of CT fracture character, RB release type and Threshold Sum had TSS values of 0.24, 0.35, and -0.12 respectively (Gauthier and Jamieson, 2008b). These results indicated that the PST was likely the better predictor of propagation propensity alone, especially given that the other tests exhibited a decrease in predictive skill when only required to predict propagation propensity.

The ECT shows the best predictive strength in all three categories based on the data collected by Simenhois and Birkeland (2009) and presented by Schweizer and Jamieson (2009). Alternatively, a small dataset collected in 2008 for this study (Table 4.8) as part of the ECT and PST comparison showed a much lower overall accuracy, TSS, and UAA at sites of confirmed initiation. An explanation for these lower scores is proposed in Section 4.4. In addition, the recent spatial variability study by Hendrikx and others (2009) found that on certain non-skier-tested slopes in which 16 to 35 ECTs were performed, about half the results were ECTP (propagated) and half were ECTN (initiated without propagating), indicating either highly variable conditions across the slope or an accuracy of around 50%. The authors suggest the latter.

Table 4.8: ECT validation results from the small dataset collected in 2008 for this study (n = 50) with the calculated prediction strengths of overall accuracy, TSS, and UAA.

PROPAGATION:		OBSERVED	
		YES	NO
PREDICTED	YES	25	4
	NO	19	2

Prediction Strength
Accuracy: 58 %
TSS: 0.23
UAA: 0.62

4.2.8 Discussion of slope-scale validation

The different measures of predictive accuracy and skill (overall frequency, TSS, UAA) used to evaluate the PST and to compare it to other tests have all been used previously (e.g. Gauthier and Jamieson 2008b; Schweizer and Jamieson, 2009), but no single method captures every aspect of measureable skill. Although the TSS and UAA are available to make *relative comparisons* when evaluating the accuracy of an unbalanced dataset, the overall frequency calculation can provide a valuable approach. Consider a prospective PST user wanting to decide if propagation will be likely if he or she initiates a weak-layer fracture. When approaching the test/profile site, he or she will not *know* if the adjacent slopes are stable, and would benefit from knowing the probability that their one test will be correct or not, independent of the real propagation propensity of the surrounding slopes. On the other hand, the overall accuracy lowers the number of each false prediction-type by calculating its frequency as a fraction of all slopes rather than as a fraction of the slope stability it failed to predict. This is where the TSS and UAA can improve the interpretation of predictive skill.

In all three measures the PST performed exceptionally well in comparison with other standard and emerging snowpack tests (Table 4.7), particularly when only propagation propensity was predicted and not initiation *and* propagation. Despite this success, a high number of false-stable predictions were observed each year, particularly in 2008. Shallow, soft slabs continued to be a source of false-stable results as Gauthier (2007, pp. 135) had observed, and day-old avalanche sites produced a substantial number of false-stable results in 2008 and 2009. Although day-old avalanche sites have been successfully used to validate other snowpack tests such as the compression test (van Herwijnen and Jamieson, 2007a) and rutschblock (Jamieson and Johnston, 1993a), Birkeland and Chabot (2006) also discussed the possibility of the snowpack subsequently strengthening after a nearby avalanche and producing false-stable results in stability tests. Perhaps in some cases where fracture initiation is still possible the energy balance required for sustained propagation may no longer exist. In other words, the 'strength' of the weak layer tested in the CT and RB may remain low at day-old avalanches and thus the layers still react, whereas the propagation propensity on the flanks may dissipate after an avalanche if the fracture toughness of the weak layer (or slab and weak layer) increases.

Many of the sites where false predictions – particularly false-stable – were observed also had correct predictions on the same layer in the same snowpit. These results are congruent with observations of slope-scale spatial variability affecting test results (e.g. Campbell and Jamieson, 2007; Hendrikx and Birkeland, 2008; Schweizer et al., 2008), and suggests that doing more than one PST can improve interpretation and reduce incorrect predictions, especially if the more conservative (propagation-likely) result is taken. The same has been shown to apply to other snowpack tests (Birkeland and Chabot, 2006; Schweizer and Jamieson, 2009; Winkler and Schweizer, 2009). It is also possible that the different test results on the same validation site-layer could be attributed to some operator error, or the mechanics of the test itself, as discussed below. Overall, however, the PST showed consistent reproducibility of results between multiple tests in the same snowpit, with only 15% of sites producing disagreement between multiple tests.

The results presented throughout Section 4.2 bring to light two important potential causes for incorrect predictions in the test: spatial variability in the snowpack, and physical differences between the mechanical process of the test and true propagation in a natural, three dimensional, undisturbed snowpack. Spatial variability of snowpack layers and test results is not addressed in detail here, although numerous studies have shown significant spatial variability in test results within a few metres of each other (e.g. Campbell and Jamieson, 2007; Hendrikx et al., 2009). The important point is that the test may not always be inaccurately assessing propagation potential when a false prediction results, since it may be accurately sampling the spatial variability that exists both across the immediate test site, and across the adjacent slope used to validate it. After all, there is another potential reason why the later-tested flanks or crown of an avalanche did not release with the avalanche: perhaps those areas were stable and *never had* high propagation propensity.

There is also the possibility that false predictions are the result of test geometry and mechanics inadequately replicating real propagation on three-dimensional slopes. For example, van Herwijnen and others (2008) and Heierli and others (2008b) chose to use longer columns to fully develop a propagating bending wave and demonstrate theoretical arguments. McClung (2009a) proposed that test-column lengths be at least three to four times the slab thickness. He supported this with observations that critical cut lengths required for propagation are a significant fraction of slab thickness and therefore the column should be much longer to observe a truly propagating fracture free from the influence of boundary conditions. Gauthier (2007) observed similar relationships between critical cut length and slab thickness in some PSTs. However, the objective of the PST is not to measure the true critical length – be it a saw cut, deficit zone, anticrack, or slowly expanding shear band – required for propagation in an undisturbed snowpack, but rather to test the general propensity for propagation to begin and the ability of the slab and weak layer to *sustain* propagation *once* it begins. In this sense, the standard PST method may not appropriately test natural initiation lengths or realistically replicate slab deformation required for self-propagating fractures (McClung 2009a), but identifies most slab and weak layer combinations in which a propagating fracture can easily begin and be sustained (Table 4.3). The only consistent exception appears to be in shallow, soft slabs where perhaps the lateral support provided by the three-dimensional snowpack on a natural slope sustains the slab and weak layer mechanical balance required for propagation, whereas in the test the soft slab breaks (SF results) more easily. McClung (2009a) also experimented with columns in which the up-slope end was left attached to the undisturbed snowpack, arguing that since a free (isolated) end attracts a crack, leaving it attached should improve testing of propagation propensity and propagation arrest. However, at present no validation data have been published to test this hypothesis.

Despite these theoretically based arguments for alternative test methods, the objectives of the standard PST method used and validated here must remain clear: to develop a *practical* test that is both efficient to perform and effective at predicting propagation propensity. The practice of guiding, forecasting, and enjoying weekend recreation are time-constricted, and therefore a test that is *efficacious* (and validated), and relatively easy to learn and perform in a short amount of time, is essential to ensure its success.

4.2.9 Summary of slope-scale validation

Predictive accuracy of the PST was fairly consistent over the three years of validation study, especially when comparing 2007 and 2009. In 2008, the accuracy was substantially lower due to a high proportion of false-stable results at day-old avalanche sites. The predictive accuracy, TSS and UAA of the PST were comparable to – and often exceeded – other commonly used snowpack tests and assessments, and the PST rarely overestimated propagation propensity. However, it produced a high number of false-stable predictions indicating low or no propagation propensity next to avalanche and whumpf sites that was higher than other common test methods. When both correct and false predictions at day-old avalanche sites were removed, PST accuracy, TSS and UAA all improved. This suggests that areas surrounding the flanks and crown of an avalanche may still contain the weak layer capable of fracturing but that propagation propensity has since dissipated, or perhaps never actually existed beyond the released avalanche. Furthermore, expecting thin soft slabs to be a source of false-stable predictions allows a more cautious interpretation of the PST or the use of an alternative test in those conditions. In order to remain practical, the PST includes some limitations, notably for thin soft slabs, but was successful in predicting slope-scale propagation propensity in most of the conditions tested during this study.

4.3 Short-scaled PST Results

During the winter of 2009, pairs of short-scaled PSTs (Section 3.6) were performed beside pairs of standard PSTs (Section 3.3) to subjectively evaluate their performance and to assess their predictive accuracy. Short-scaled PSTs were tested to determine if scaling the column length below 1 m would reduce the number of false-stable predictions at shallow, soft-slab avalanches observed in the standard PST validation dataset. A subjective review of qualitative field observations from using the short-scaled PST is presented in Section 4.3.2, and the validation results of the alternative method are presented in Section 4.3.3.

4.3.1 Data

In total, 231 short-scaled PSTs were performed of which 73 coincided with the validation study, enabling the predictive accuracy of the short-scaled method to be evaluated. Since there was potential from the onset for the short-scaled method to replace the standard 1 m column if validation showed improved results, a full range of depths (and therefore column lengths) between 10 cm and 93 cm was tested (Figure 4.4). This was done to assess whether the short-scaled method was effective and accurate over the full range up to 1 m, and not just in the targeted shallow layers. Raw data are presented in Appendix B.

4.3.2 Qualitative observations

Throughout the season, operators subjectively evaluated the effectiveness of the short-scaled method. For longer columns with thicker slabs (i.e. > 50 cm), the test performed similarly to the standard 1 m long method. However, testing in the field showed that short-scaled PSTs less than 30 cm long (typically) were difficult to perform properly. The slab would commonly pivot at the saw around mid-point in the column rather than allowing the fracture to propagate, particularly in soft storm-snow slabs overlying persistent weak layers. In other cases, slight upward force inadvertently generated by the

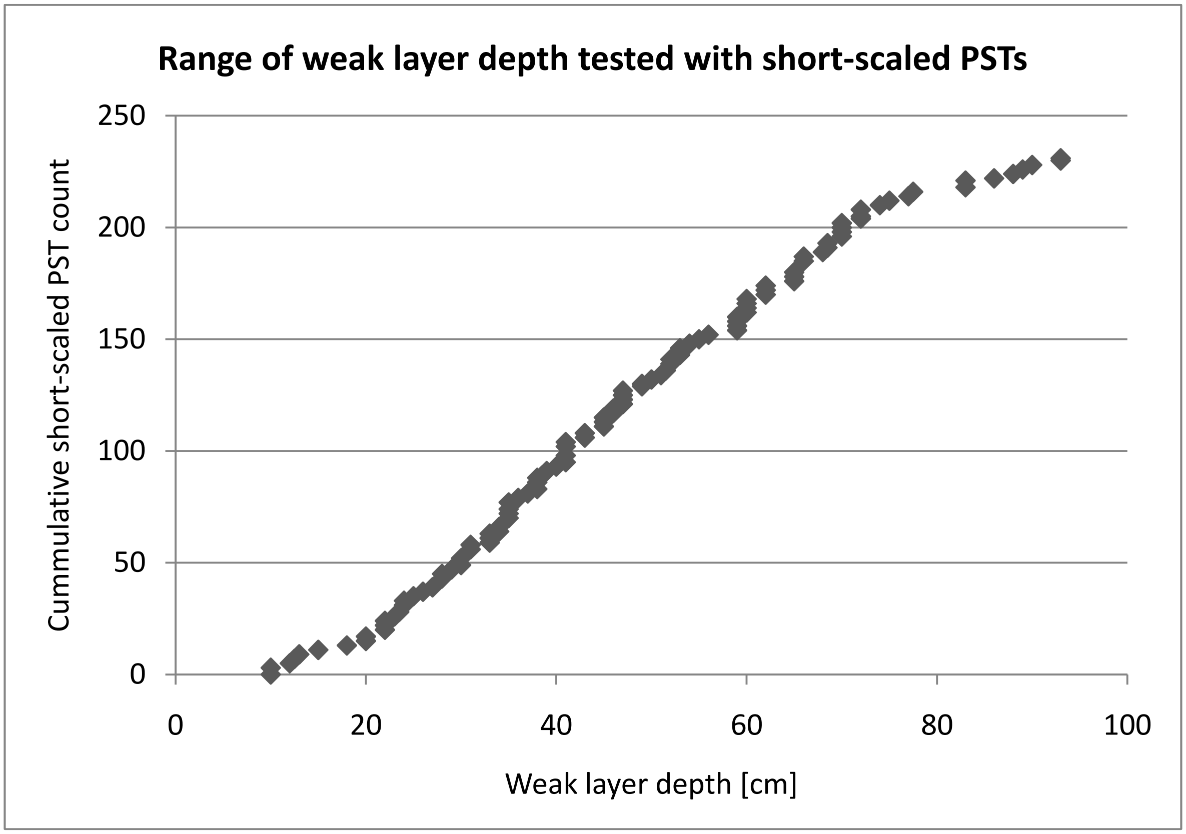

Figure 4.4: Tested weak layer depth plotted against the cumulative count of short-scaled PSTs (n = 231), showing the full range (10 cm to 93 cm) of slab thickness and equivalent column length tested with the short-scaled PST in the 2009 season.

operator while dragging the saw would lift the soft slab off the underlying part of the column. Although some test columns less than 30 cm long appeared to propagate the fracture rather than pivot, this was hard to verify. In addition, a large percentage of these results had near-50% cuts, which are interpreted literally with the < 50%-cut rule, but potentially with less confidence. The short-scaled PST was rarely performed in stiff, thin slabs (i.e. wind slabs), but showed improved ease-of-use and performance in those cases. Short-scaled PST results are plotted against weak layer depth (equivalent to column length) in Figure 4.5 showing the large number of shallow slabs that pivoted and the high number of results with near-50% cuts.

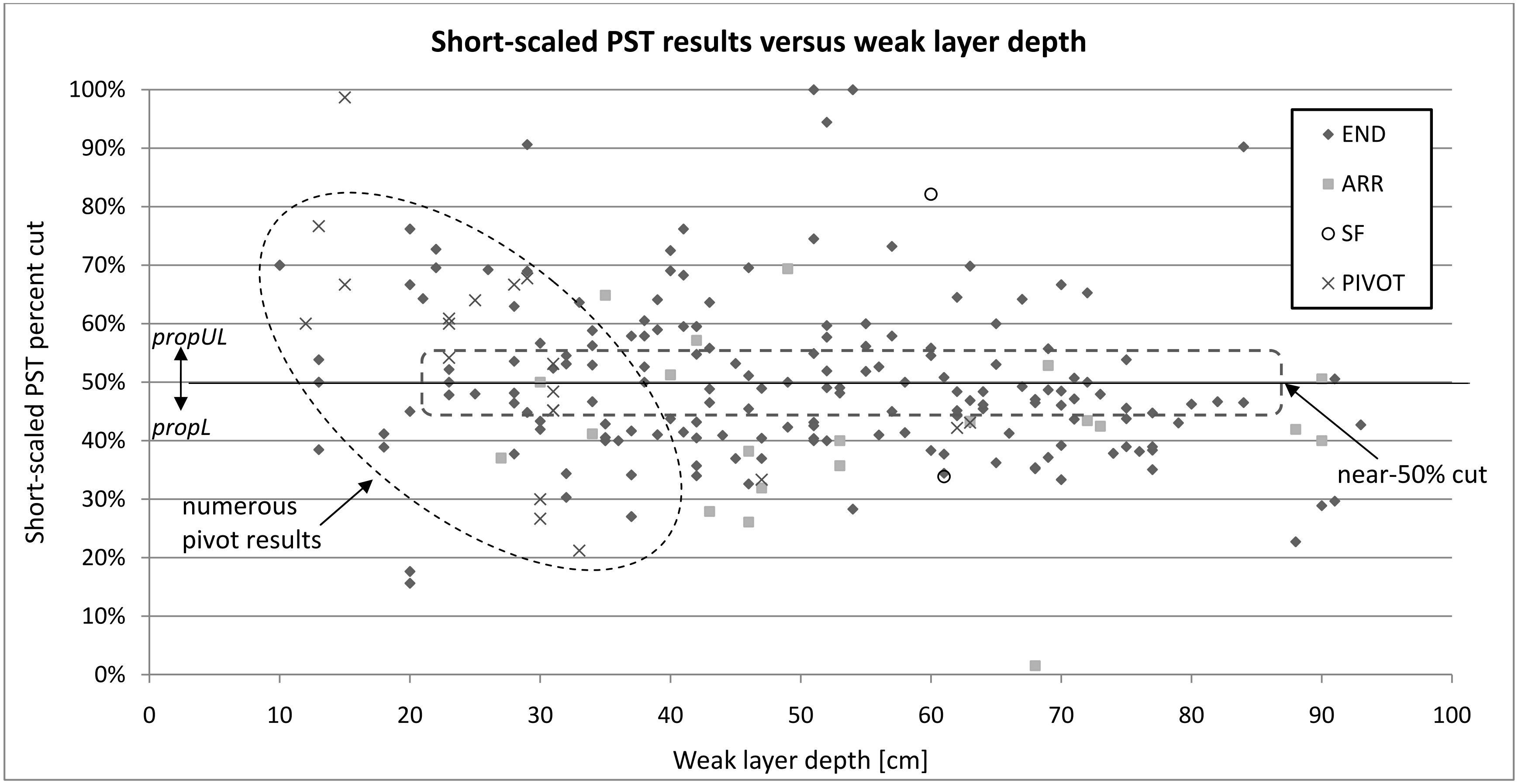

Figure 4.5: Short-scaled PST results plotted against weak layer depth, which is equivalent to column length by the short-scaled method. Four symbols differentiate the standard results of END, SF and ARR, as well as slabs that pivoted around the saw during the test. Numerous PIVOT results are evident in shallow layers, although many cases may not have been recorded. Additionally, the large number of tests with cut lengths near 50% is evident. END is used for cases where 100% of the column was cut without propagation.

4.3.3 Slope-scale validation

Despite the often-problematic performance in shallow, soft slabs, the short-scaled PST was performed at validation sites alongside the standard PST to test its predictive accuracy. A total of 73 short-scaled PSTs were performed on 31 validation site-layers in 2009, of which 32 were on layers showing confirmed propagation and 41 were on layers with confirmed initiation without propagation. The results are shown in Table 4.9. The short-scaled PST results were interpreted with the same rule as the standard PST; however, due to the large percentage of tests with cut lengths close to 50% (plus or minus 2 cm) which are generally interpreted with less confidence these results are separated in Table 4.9.

Table 4.9: Validation results of short-scaled PSTs next to sites of confirmed propagation or initiation without propagation. The near-50% cut results are separated initially due to lower confidence interpreting their results, and then interpreted literally based on the rule of < 50% cut predicting propagation-likely. Slabs that pivoted were kept separate in both cases.

Analysis:	**50% cut +/- 2 cm results separated**		**50% cut +/- 2 cm interpreted based on PropL< 50%≤PropUL**	
Short-scaled PSTs	**Total**	**%**	**Total**	**%**
Correct Predictions	33	45%	39	53%
False-Stable	6	8%	10	14%
False-Unstable	8	11%	11	15%
	↓ **separate, un-interpreted results** ↓			
50% cut +/- 2cm	13	18%	~	~
Slab Pivoted	13	18%	13	18%
Total	73	100%	73	100%

Even when interpreted literally based on less than 50% cut predicting propagation likely, the frequency of correct predictions only increased from 45% to 53%, while at the same time false-stable results and false-unstable results increased by six percentage points and four percentage points, respectively. In addition, slabs that pivoted are included in the table to demonstrate the frequency at which they were observed.

At two sites of confirmed propagation in 2009 – each with shallow layers less than 35 cm deep – the two short-scaled PSTs at each site correctly predicted propagation where the two standard PSTs failed. Despite this small reduction in false-stable results in a few shallow, soft-slab cases, scaling the PST below 1 m did not improve the predictive accuracy of the test overall and in fact made it more challenging to perform on those types of slabs.

4.3.4 Summary

A qualitative assessment of the short-scaled PST method in the field and an examination of validated short-scaled PST results were conducted in the winter of 2009. The results revealed that, despite two cases of improved predictions in soft, shallow slabs, the short-scaled method neither improved the predictive accuracy of the PST overall nor proved functional in the field in many of the same shallow, soft slabs that were targeted for improvement. Since the primary purpose of scaling the PST below 1 m was to potentially reduce false-stable predictions in the same shallow, soft slabs it performed poorly in, the short-scaled PST method cannot be recommend. Instead, the standard 1 m+ column length method has been shown to be the more reliable and accurate predictor of propagation on adjacent slopes, and is recommended along with the recording standard presented in Section 3.3.

4.4 Practical Results of the PST and ECT: Strengths and Limitations

The PST and ECT are recent additions to the forecast and snowpack assessment toolset, and as such, practitioners are still learning their methods and subjectively evaluating each test, which will ultimately establish their credibility and define their niche in the toolset. Given that most practitioners have little time to experiment with new tests, the objective of this comparison is to report on the substantial data collected for this study in 2008 in order to present the strengths and limitations to each test method individually and in comparison with each other. The objective in this section is not to promote one test method over the other as the 'better overall predictor of propagation' but to determine particular conditions under which one test method may be more efficacious than the other, specifically in relation to slab and weak layer combinations with varying degrees of propagation propensity.

Initial observations in the field had hinted at snowpack characteristics, particularly slab thickness (and hardness), which influenced PST and ECT results. To assess the effect that changing slab thickness had on test results, PST and ECT results were compared separately against weak layer depth in Section 4.4.2 and 4.4.3 respectively, and directly in Section 4.4.4.

4.4.1 Data

A total of 246 ECT results were recorded in 183 individual tests performed in 91 snowpits on 88 field days in 2008. Since the ECT is loaded by tapping on a shovel blade on the surface, multiple layers frequently fractured in the same test column. Often, fractures in shallow storm layers would initiate, as commonly seen in the compression test, without any indication of propagation. Since the objective in performing the ECT was to compare results to the PST, ECT results were often only recorded for those layers also tested in the PST. These were primarily the persistent weak layers listed in Section 3.2.1. For the same reason, particularly in March 2008 when the reactive 23 February and 9 March surface hoar layers were close to each other in the snowpack, multiple comparisons could be made from within the same PST and ECT columns. In other words, two layers could be saw-cut in the same PST column which would both usually react in the ECT as well. Because of this, the 91 snowpits produced 99 *comparison groups*, where comparison groups are defined as all PSTs and ECT results on a specific weak layer within the same snowpit. For the separate analysis of test results versus weak layer depth, some PST and ECT results were omitted if weak layer depth had not been recorded, or if uncertainty existed about the results.

4.4.2 Effect of slab characteristics on PST results

The results of 365 PSTs from 2008 are plotted against increasing weak layer depth in Figure 4.6. The percentage of the PST column cut at the onset of propagation was used as the ordinate measure of test results, and different symbols distinguish the propagation result that followed (*ARR/SF/END*). The results of 371 PSTs from 2009 are also plotted against weak layer depth in Figure 4.7 for comparison. The seasonal PST data are separated to show repeatable trends in independent datasets and to reduce clutter in the plots.

In the winters of 2008 and 2009, researchers performing the PST in the field had observed that shallow weak layers with soft overlying slabs often resulted in slab-fracture results, which coincided with numerous false-stable predictions in the validation data (Section 4.2). Field workers had also observed that a minimum slab thickness, and perhaps hardness, was required to facilitate propagation to the end in test columns, and that the PST could indicate high propagation propensity (*propL)* in layers two or more metres below the snow surface. Data presented in Figures 4.6 and 4.7 reflect these field observations.

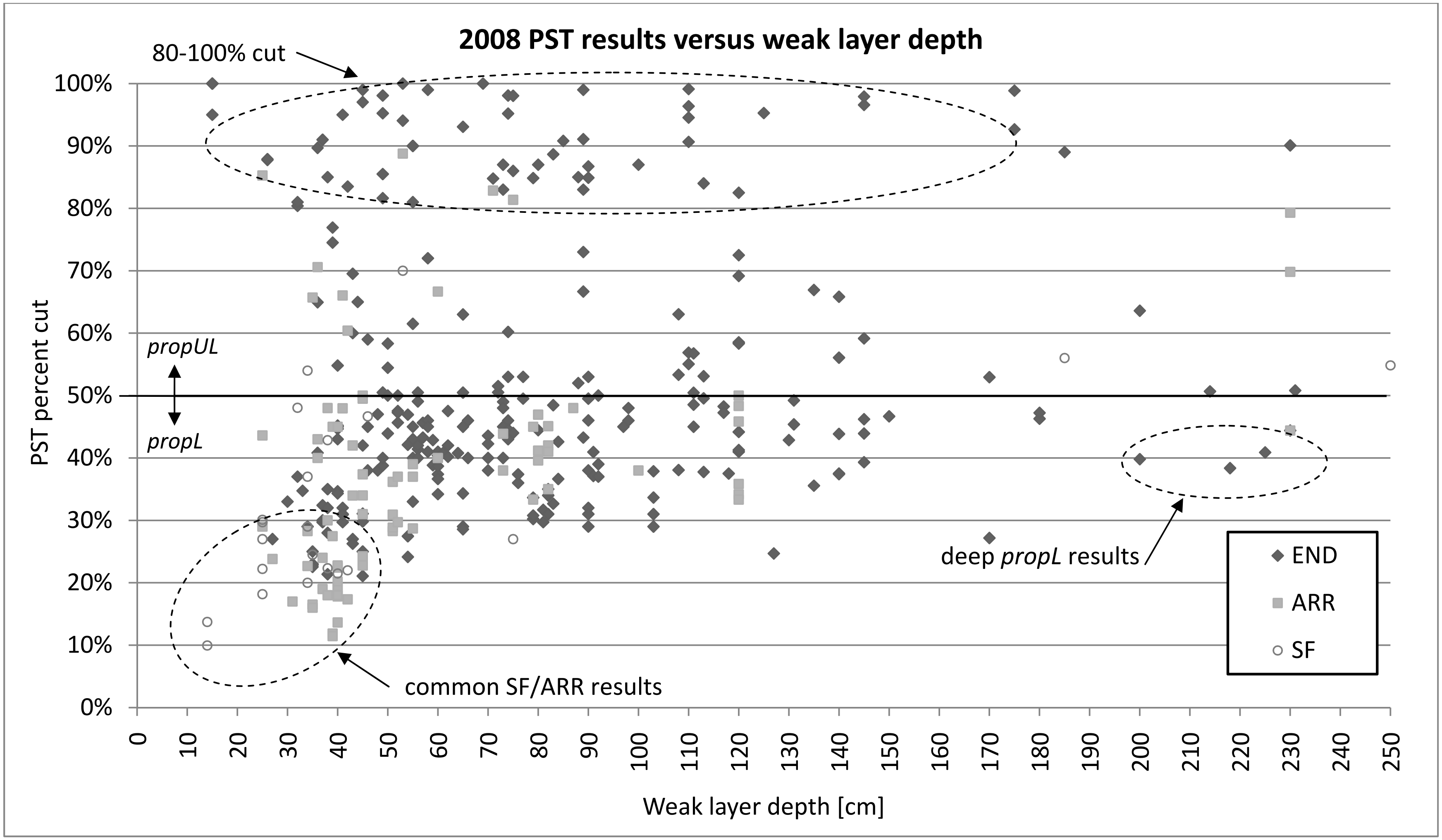

Figure 4.6: PST results from 2008 versus weak layer depth (n = 365). Percent cut at the onset of propagation is used as the ordinal measure with three symbols to distinguish between the three possible results of propagation to the end (END), self-arrest (ARR) or slab fracture (SF). Only END results up to 50% cut indicate propagation is likely (propL) on adjacent slopes.

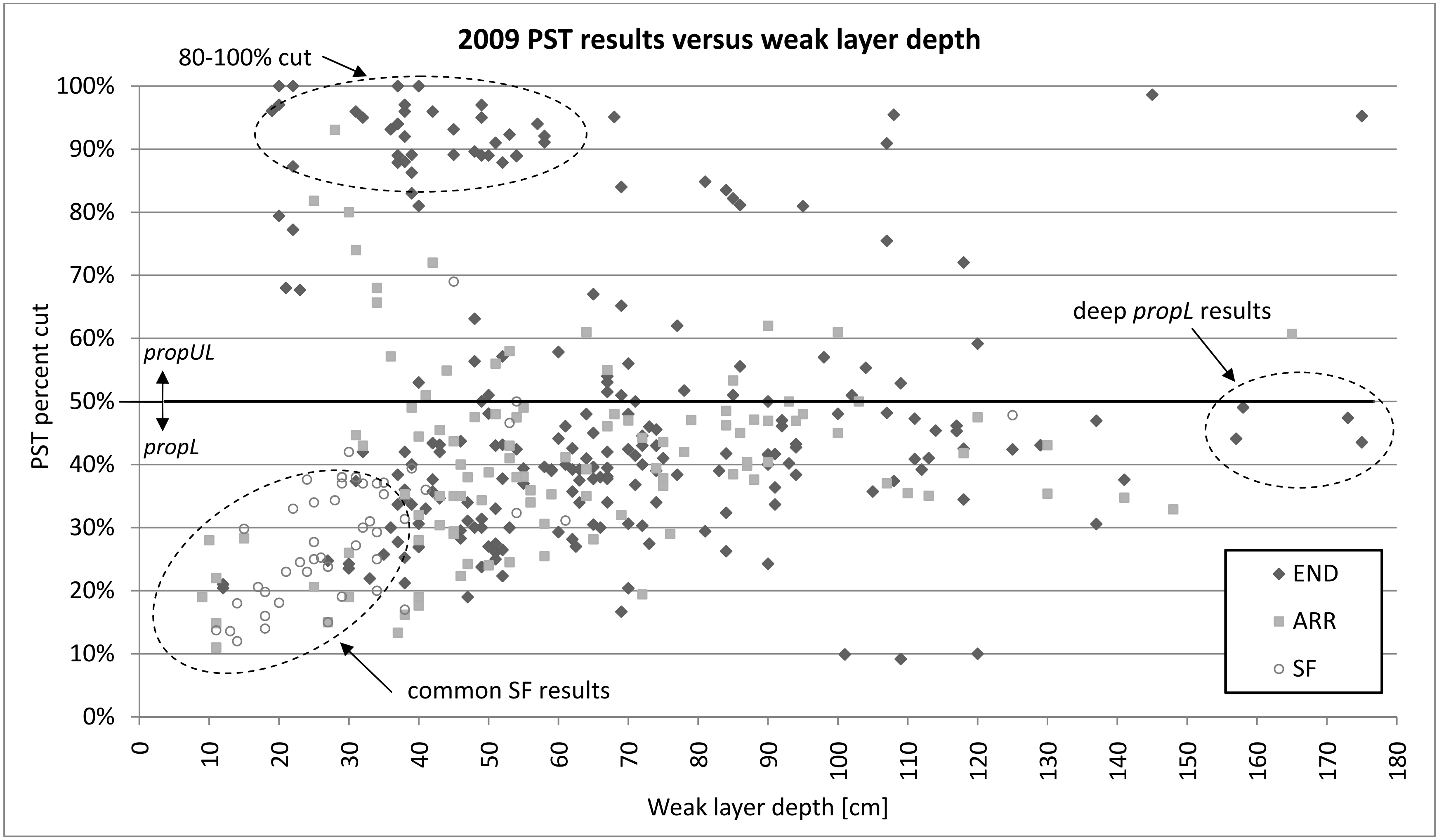

Figure 4.7: PST results from 2009 versus weak layer depth (n = 371). Percent cut at the onset of propagation is used as the ordinal measure with three symbols to distinguish between the three possible results of propagation to the end (END), self-arrest (ARR) or slab fracture (SF). Only END results up to 50% cut indicate propagation is likely (propL) on adjacent slopes.

Both figures show an apparent trend of increasing percent-cut with depth for initiating propagation in shallower layers up to around 50 or 60 cm. Beyond that point the data show more scatter. Gauthier (2007, pp. 114) had found significant correlations with slab thickness and hardness when seeking *critical cut lengths* in various-sized columns, although that differs from the effect on cut-*percent* shown here. In the dataset analyzed here, percent-cut is equivalent to cut length up to 100 cm, after which length-scaling caused percent cut and column length to diverge.

Many of the PSTs in shallow layers ended in SF or short propagations to ARR, particularly in 2009 where more shallow layers were tested. In fact, SF results occurred most commonly in thin slabs between 10 cm and 40 cm thick and became infrequent beyond that. The hand-hardness of the slab in these cases was generally fist (F) to fist-plus (F+) indicating soft, low-density slabs. Short propagations, whether arresting within the layer or going to the end after well over half of the column had been cut, are widely dispersed throughout the datasets in both figures.

The shallowest layer with both an END result and a cut less than 50% occurred under a slab 9 cm thick. However, with that exception, this combination of results (*propL)* did not begin to occur until the slab thickness exceeded 30 cm. At greater depths, results propagating to the end predominantly required cuts between 30% and 60%, or greater than 80%. The latter observation might be the result of the free up-slope end of the column attracting a fracture, as McClung (2009a) proposed as the reason for not isolating that end in his alternative test method. In other words, when propagation propensity in the tested snowpack is very low, one would expect to be able to cut right to the end of the column without any onset of propagation, which would explain why very few propagating fractures started between 60%- and 80%-cut when the PST indicated *propUL*. However, a large number of observed results *did* propagate a fracture to the end after 80% to 100% of the column had been cut, perhaps because the free end attracted the fracture (McClung 2009a).

It is also important to note the number of tests that predicted high propagation propensity in weak layers deeper than 100 cm. *PropL* results were observed in tests up to 230 cm deep in 2008 and up to 175 cm deep in 2009, and Gauthier (2007, pp. 195) observed a *propL* result in a slab 285 cm thick in a previous season. Of particular interest are the two tests between 120 and 180 cm deep in 2008 that propagated to end after less than 30% of the column had been cut; and the two tests in 2009 between 100 and 120 cm deep that propagated to the end with only about 10% cut. All had a 20-30 cm thick pencil-hard layer immediately overlying the weak layer. Some PSTs also resulted in ARR or SF in deeply buried weak layers indicating that the PST appears capable of differentiating between high and low propagation propensity in deep slabs. Validation of these deep-slab PST results is presented in Section 4.5.

4.4.3 Effect of slab characteristics on ECT results

Test scores from 242 ECT results are plotted against weak layer depth in Figure 4.8. No ECTPV (initiation and propagation during column isolation) results were observed. ECTX results were given a score of 35 in order to be plotted against weak layer depth. Results on layers that were not comparatively tested with the PST are included here.

In terms of test score, ECT results in Figure 4.8 show a clear dependence on depth, with thicker slabs requiring more taps to initiate and potentially propagate fractures. Despite this, the different results of ECTN and ECTP do not show a dependence on depth in slabs between 30 cm and 70 cm deep. This reveals that score was more dependent on depth through this range than was the indication of propagation once a fracture was initiated. In weak layers less than 30 cm deep, fracture initiation was common without propagation; and for weak layers deeper than 70 cm, fractures were occasionally initiated, of which only some propagated.

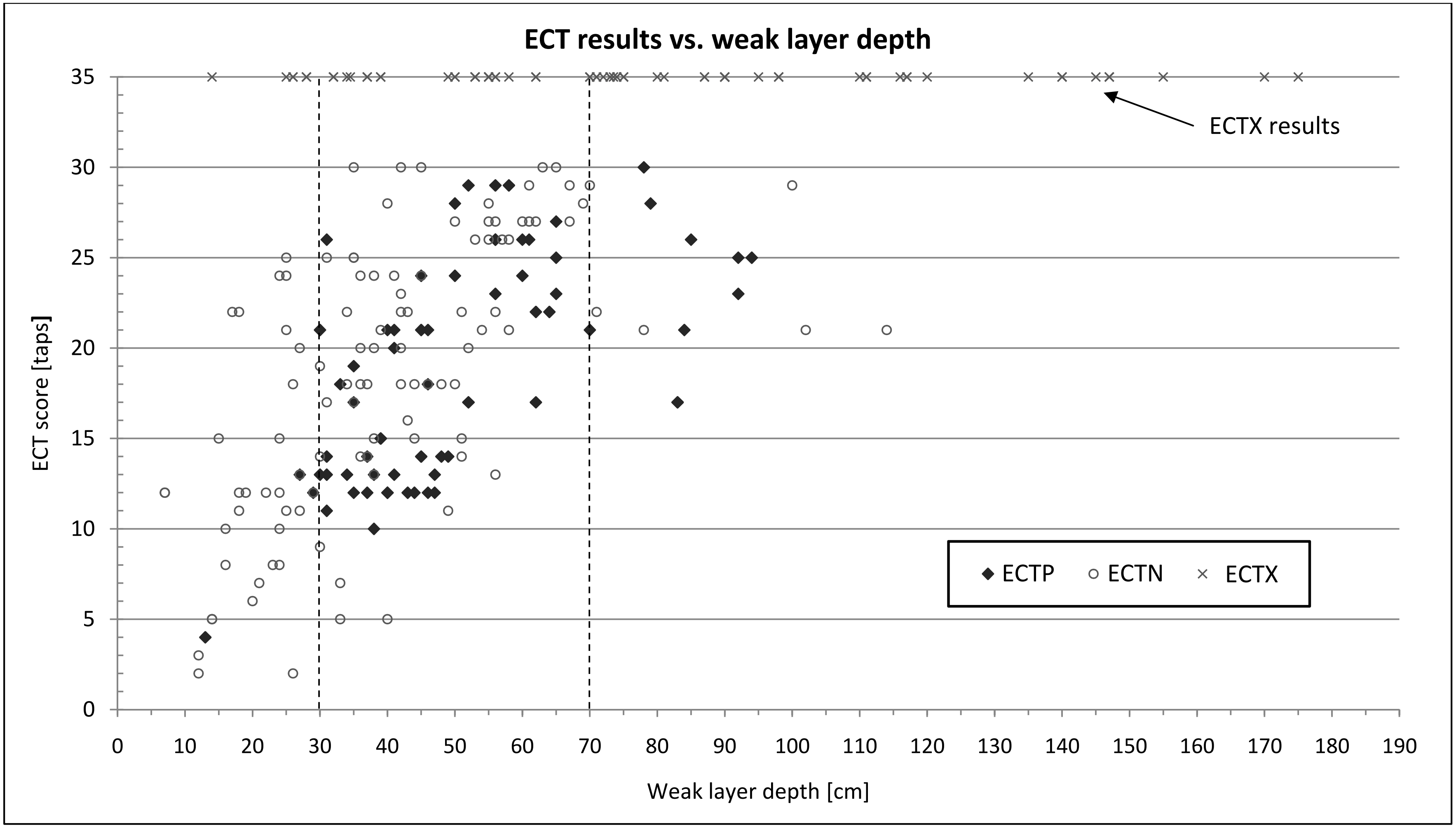

Figure 4.8: ECT results from 2008 versus weak layer depth (n = 242). ECT score (number of taps) is used as the ordinal measure with three symbols to distinguish between the three types of observed results (ECTP, ECTN, ECTX). ECTX results were given a score of 35 in order to be plotted.

Since the number of ECT taps is implied to be an indicator of the ease of fracture initiation (Simenhois and Birkeland, 2006), it was expected that as layer depth increased, an increased number of taps was required to initiate (and potentially propagate) fractures. From Figure 4.8 it can be seen that 85% of ECTP results occurred when the slab thickness was between 27 cm and 71 cm, when only 65% of ECT-tested slabs were in this range. An additional 13% fell between 71 cm and 92 cm (compared to 10% of all tests), and only one incidence of ECTP occurred in a slab thinner than 27 cm even though 14% of tests were in such slabs. This particular case occurred on a storm snow interface 13 cm deep with an overlying F to F+ slab.

In slabs thicker than 92 cm, a few fractures initiated up to 114 cm deep – all without propagation, but most tests had no result. The data in Figure 4.43 show what researchers had subjectively observed in the field: that the ECT indicated propagation when the conditions favourable to initiation were present, but could not indicate propagation propensity once the weak layer was too deep for a fracture to be initiated in the standard 30 loading steps (taps).

4.4.4 Direct comparison of PST and ECT results

Direct comparisons of PST and ECT results are based on the interpreted test results of *propL* and *propUL*. Agreement between interpreted results from different test methods applied within the same snowpit is desirable from a forecasting perspective as they give confidence to a snowpack assessment. The results presented here compare the frequency at which tests were in agreement, but have not been validated to determine which test(s) were *correct* about propagation propensity. Only ECT results on layers that were also tested with the PST are included here. A small validated and comparable dataset exists as presented separately in Section 4.2 and together in Section 4.4.5.

Within each of the 99 comparison groups defined in Section 4.4.1 the first two ECTs and first two PSTs performed were compared based on *propL* and *propUL* (Table 4.10a). All four tests agreed on propagation propensity in 51% of comparison groups. In an additional 25% of groups three out of four tests agreed. In the remaining 24 groups (24%) both PSTs agreed with each other but were in disagreement with both ECTs. There were no cases where one PST and one ECT indicated *propL* while the other pair indicated *propUL.*

Table 4.10: Comparative results of two PSTs and two ECTs within the same pit on the same weak layer based on the interpreted results indicating propagation was likely (propL) or propagation was unlikely (propUL). The results from all comparison groups (a), and only those with weak layers less than 70 cm deep (b), are given. Cases where both PSTs gave opposite results to both ECTs are separated into their two possible scenarios.

a)	**All four agree**	**Three of four agree**	**PST opposes ECT**	**One of each agrees**
All comparison groups (*n* = 99)	50 (51%)	25 (25%)	24 (24%)	0 (0%)
	PST *propL* vs. ECT *propUL:*		18	
	ECT *propL* vs. PST *propUL:*		6	

b)

Comparison groups with targeted weak layers < 70 cm deep (*n* = 46)	29 (63%)	11 (24%)	6 (13%)	0 (0%)
	PST *propL* vs. ECT *propUL:*		2	
	ECT *propL* vs. PST *propUL:*		4	

Of the 24 groups of conflicting results between test methods, 18 (75%) involved the PST indicating *propL* while the ECT indicated *propUL*. In all but two of those groups the weak layer of interest was 70 cm or deeper; and only twice did a fracture initiate within the standard loading steps of an ECT – the rest were ECTX. The remaining six test groups in which the PST and ECT disagreed saw the ECT indicating *propL* while the PST indicated *propUL*. One of these groups was possibly due to the previous collapse of the weak layer during a nearby avalanche; while in another, both PST results were just over 50% cut. Another three of these groups occurred over three consecutive days of testing the

January 26 surface hoar layer, in which the slab was generally finger-hard (1F) settled snow overlying four-finger-hard (4F) surface hoar with grain sizes of 2-6 mm.

When only slabs shallower than 70 cm were considered (Table 4.10b), three or four of the four tests within a comparison group agreed 87% of the time and the number of ECT *propUL* versus PST *propL* comparisons was reduced from 18% to 4%. This confirmed that where the ECT no longer indicated *propL* in deeply buried weak layers, the PST showed that propagation in those layers could sometimes still be high. In most of these cases, there was no result in the ECT. In other words, the applied surface loading did not penetrate deeply enough into the snowpack to initiate a fracture in the ECT, which *may* have propagated *if* initiated. This is an indication that at some limiting depth the ECT will not show the true propagation propensity of a layer if a fracture cannot be initiated through surface loading in the standard 30 taps.

When a one-to-many comparative analysis was performed on individual tests within the same comparison group, similar results emerged (Table 4.11). For this analysis each PST was compared to each ECT in the same group, generating *m x n* comparisons, where *m* and *n* are the respective numbers of PSTs and ECTs within a group.

Table 4.11: One-to-many comparisons of each test within the same pit on the same layer, generating m x n comparisons, where m and n are the number of PSTs and ECTs, respectively. n = 432.

All tested layers		**PST**	
		propL	**propUL**
ECT	**propL**	97 (23%)	36 (8%)
	propUL	129 (30%)	170 (39%)

A total of 432 direct comparisons were made of which 62% (267) agreed on whether propagation propensity was likely or unlikely. As before, a larger number of PSTs indicated

PropL while the ECT indicated *PropUL* (30%), primarily in deep weak layers where fractures did not initiate in the ECT. In only 8% of comparisons the ECT indicated *propL* while the PST indicated *propUL.*

When the same comparative analysis was performed on only tests in layers shallower than 70 cm (Table 4.12), the number of agreements improved from 62% to 74%, and the number of times the ECT indicated *propUL* next to a PST indicating *propL* decreased from 30% to 13%.

Table 4.12: One-to-many comparisons of each test within the same pit on the same layer, where only tests on layers shallower than 70 cm were included. n = 238.

Layers < 70 cm deep		**PST**	
		propL	**propUL**
ECT	**propL**	79 (33%)	30 (13%)
	propUL	31 (13%)	98 (41%)

These results further confirmed the subjective field assessment that once a weak layer was buried deeper than 70 cm in the snowpack at the test site the ECT no longer reliably initiated fractures in that layer which, when tested with an adjacent PST, frequently showed high propagation propensity.

To further compare tests, the first (or second) PST result from a pit was plotted against the first (or second) ECT result from the same pit (Figure 4.9). This analysis used a one-to-one comparison of PST to ECT results rather than the one-to-many comparison used previously. In other words, the first PST and ECT from a test group were compared, as were the second PST and ECT, but the first PST and second ECT were not. Although this reduced the number of comparisons, the trends in the data were still apparent. To plot results, the percentage of the PST column cut at the onset of propagation is the ordinate measure while the ECT score is shown on the abscissa. Four series were used to distinguish the interpreted results and the letter L indicates the test predicted propagation (equivalent to *propL)*, while the letter U indicates propagation was *not* predicted (*propUL).* The first letter in each series refers to the ECT result and the second refers to the PST result.

Most comparisons that agreed on propagation being likely (LL) occurred within the range of 20%- to 40%-cut in the PST and scores between 10 and 30 in the ECT. This reaffirms that *PropL* PST results can show consistent cut percentage whilst propagation in the ECT depends on initiation ease and thus the full range of scores is represented. No comparisons fall near the origin, where PST cuts are short (< 30%) and ECT score is low (< 10). This is partially due to limited testing with the PST in shallow layers, but may also be related to thicker, stiffer slabs favouring propagation (e.g. van Herwijnen and Jamieson, 2007). Although numerous ECT results with scores less than 10 are plotted in Figure 4.9, they were primarily ECTN results in storm layers that were not tested with an adjacent PST and therefore are not analyzed in this section.

In addition, most of the comparisons in which the ECT indicated *propUL* and the PST indicated *propL* (UL) required cuts between 30% and 50% of the column length and generally had higher ECT scores than the LL comparisons. This coincides with the observations presented in Section 4.4.2 showing deepening weak layers required longer PST percent cuts (likely only to a critical length as in Gauthier, 2007), and were more difficult to initiate in the ECT. Many of the UL results in the 30%- to 50%-cut range are clustered on the far edge with ECT scores of 35. These are primarily the deep weak layers that still showed high propagation propensity in the PST but could not be initiated within the 30 taps in the ECT. The few comparisons in which the ECT indicated *propL* and the PST indicated *propUL* (LU) are dispersed throughout the range of ECT score and PST percent-cut.

Using the same 203 one-to-one comparisons of an ECT next to a PST, the percentage of times the two different test methods agreed was determined. This is valuable since a

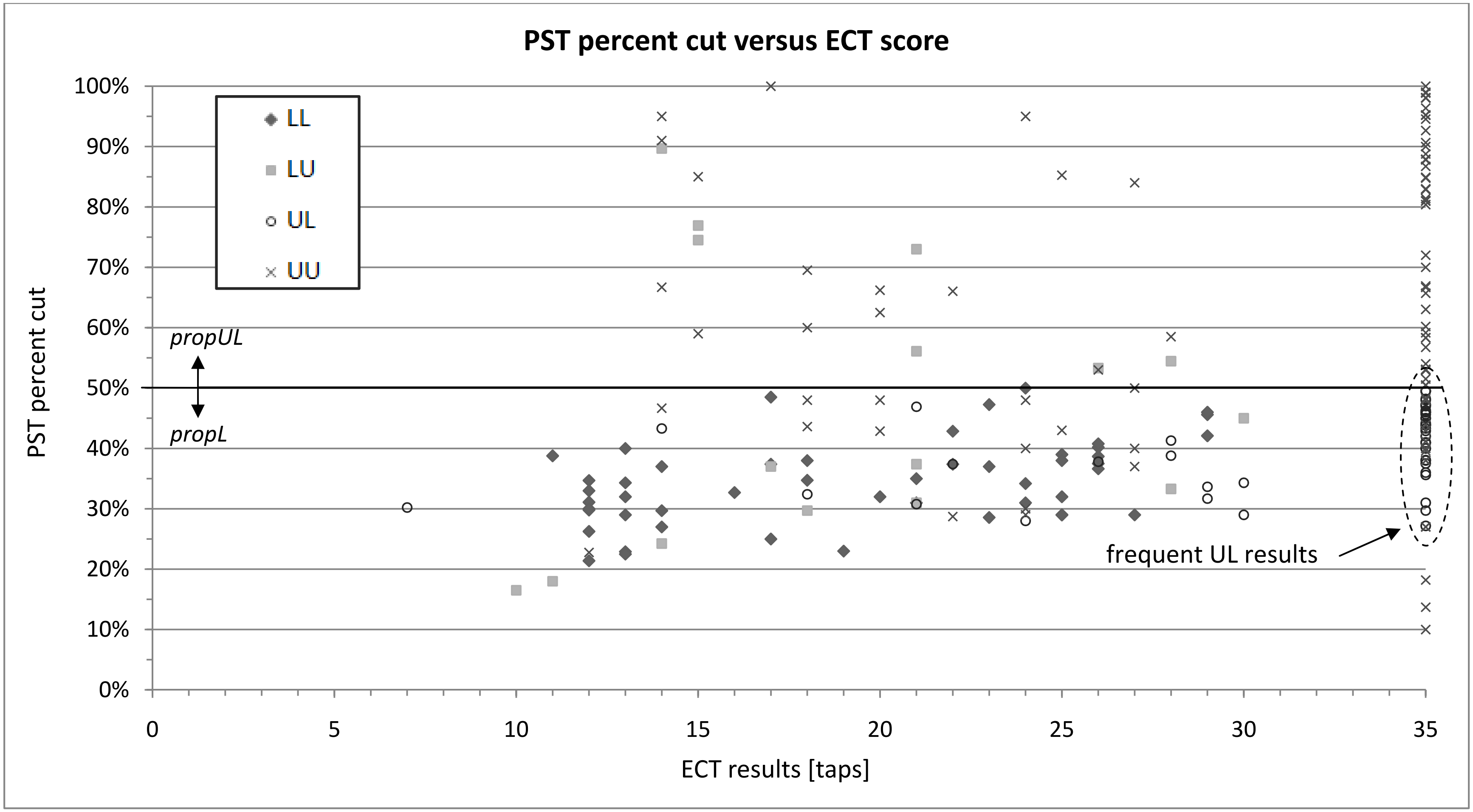

Figure 4.9: PST results (percent of column cut) plotted against ECT score (number of taps) for one-to-one comparisons of opposing pairs of tests in each test group (n = 203). In the legend, the first letter refers to the ECT result and the second refers to the PST. The data series 'LL' indicate that both tests suggested propagation was likely. 'LU' indicates the ECT suggested propL and the PST suggested propUL, while 'UL' indicates the opposite case. Finally, 'UU' indicates both tests suggested propagation was unlikely.

practitioner or recreationist *might* realistically do one of each test in their assessment and rarely more. The results are shown in Table 4.13 and reveal that pairs of tests agree 67% of the time and that the ECT does not indicate *propL* in half of the cases the PST does. Of those 50 ECTs that indicated *propUL* when the PST indicated *propL,* 14 were ECTN and 36 were ECTX. Again, these results are not validated next to observations of propagation or initiation without propagation in the field.

Table 4.13: One-to-one comparison of ECT and PST pairs from within the same pit on the same layer (n = 203).

Compared pairs:		**PST**	
		propL	**propUL**
ECT	**propL**	49 (24%)	16 (8%)
	propUL	50 (25%)	88 (43%)

4.4.5 Validated comparison of PST and ECT results

A small dataset of validated, comparable PSTs and ECTs was collected in 2008 as a product of the separate PST validation study and the ECT and PST comparative study. On 24 of the 28 validation site-layers in 2008, at least two ECTs were performed beside at least two PSTs and the prediction 'scenarios' from the first two of each test are compared in Table 4.14. In nine cases (38%) all four tests made correct predictions of propagation propensity. Conversely, in five cases (21%) all four tests *incorrectly* predicted propagation propensity. Four of these cases were at day-old avalanche sites previously shown to generate a high number of false-stable predictions with the PST alone (Section 4.2.4). In five of the 24 site-layers, three of four tests correctly predicted propagation propensity, usually with two correct PSTs and one correct ECT.

Table 4.14: The predictions of two PSTs and two ECTs at each of 24 validation site-layers in 2008 are compared. The number of times each 'scenario' occurred is given, along with any associated notes of interest. All but three of the site-layers showed confirmed propagation nearby.

Prediction Scenario		**Notes:**
All four are correct:	9 (38%)	Seven of confirmed prop., two of init. without prop.
All four are incorrect:	5 (21%)	Four of five cases were at day-old avalanche sites
Three of four correct:	5 (21%)	In four of five cases, both PSTs and one ECT was correct
Three of four incorrect:	3 (13%)	One ECT was correct in two cases, one PST in the other
PSTs correct/ECTs incorrect	1 (4%)	Layer was buried 79 cm deep.
ECTs correct/PSTs incorrect:	1 (4%)	Day-old avalanche site. Layer was 39 cm deep.
TOTAL	24	

Table 4.15 shows the results of one-to-one comparisons of tests (as described in the previous section) on the 26 validation site-layers on which at least one ECT and one PST were performed, of which 23 were site-layers of confirmed propagation and three were site-layers of confirmed initiation *without* propagation. Surprisingly, only 43% of pairs both correctly predicted *propL* and a concerning 27% both falsely predicted *propUL* next to avalanches or whumpfs, although nine of the twelve pairs were at day-old avalanche sites.

Table 4.15: One-to-one comparisons of PST and ECT results separated by (a) site-layers of confirmed propagation; and (b) site-layers of confirmed initiation without propagation.

a) confirmed propagation

	n= 44	**PST**	
		propL	**propUL**
ECT	**propL**	19 (43%)	6 (14%)
	propUL	7 (16%)	12 (27%)

b) confirmed initiation w/o propagation

	n = 5	**PST**	
		propL	**propUL**
ECT	**propL**	1	1
	propUL	0	3

An approximately equal number of PSTs and ECTs falsely predicted *propUL* when the alternate test method in the pair was correct. Not enough pairs at sites of initiation without propagation exist to develop significant conclusions, although three of five pairs accurately predicted *propUL*.

Five validation site-layers with at least one comparable PST and ECT were between 70 and 90 cm deep. In two of those cases, the PST correctly predicted *propL* and the ECT did not. A fracture only initiated in one of four ECTs on these two site-layers, but the other three showed no result (ECTX). The other three cases were at day-old avalanche sites in which all six PSTs gave false-stable predictions and five of six ECTs gave false-stable results. Only one ECT propagated after 21 taps and the other five were ECTX.

4.4.6 Discussion

Results from 2008 show that the PST and ECT generally agree upon predictions of propagation propensity in slab and weak layer combinations up to approximately 70 cm deep, after which fractures will sometimes initiate and less frequently propagate in the ECT. Beyond this depth the ECT was inconsistent in predicting propagation while the PST indicated propagation propensity in weak layers up to the tested depth of 250 cm. The wide range of slab thickness tested in this study might explain why the ECT accuracy and skill scores presented in Table 4.8 are lower than those calculated by Simenhois and Birkeland (2009) (Table 4.8), who may have sampled a shallower range of slabs. Although skier triggering (and hence validation) becomes increasingly rare at these greater depths, the PST appears better suited to identify cases where high propagation propensity still exists in a layer which may be triggered from a thin spot in the snowpack nearby. This potentially reduces the dependence on site selection for tests.

The ECT appears to have the capacity to test both fracture initiation and propagation in a single snowpack test, and has been shown here to give consistent results in weak layers buried beneath slabs 30 cm to 70 cm thick, and occasionally up to 95 cm. Beyond 95 cm, the stresses transferred from the shovel during loading appear to rarely extend deep enough to initiate fracture in the weak layer. For a recreationist attempting to assess slope stability in an unfamiliar snowpack with multiple buried weak layers, the ECT can be performed quickly and easily and seems to provide more information about propagation potential than a standard compression test.

The PST, on the other hand, requires the user to pre-identify weak layers of interest within the snowpack for testing, and makes no indication of the ease at which fractures may be triggered from the surface. This could be problematic if the user is unfamiliar with the snowpack in the area or uncertain as to which layer(s) may potentially fracture and propagate. In most cases, quick examination of a snowpit wall reveals any weak layers of interest, particularly facet layers or well-preserved surface hoar layers which combined can produce approximately 95% of persistent avalanche activity, including most large avalanches (Hägeli and McClung, 2003). In other cases, a compression test or deep tap test performed first can identify layers prone to fracture initiation, which can then be tested for propagation potential with an adjacent PST. One *advantage* of having to pre-select a weak layer for testing is that the specificity of the test requires the user to think critically about the snowpack and to identify, test, and closely observe the critical layers which are likely to play a role in avalanches. This also enables forecasters or experienced recreationists to track the propagation propensity of specific layers over time, even to great depths in the snowpack.

Neither test method frequently indicated propagation in slabs thinner than 30 cm, which appears to indicate that propagation propensity is typically low in those conditions, at least in the average Columbia Mountain snowpack. Schweizer and Jamieson (2007) found similar results in their development of the threshold sum approach to stability assessment, where they identified slabs ≥ 24 cm thick in their dataset to have sufficient stiffness to facilitate fracture propagation. It may also be the result of test geometry in the PST and ECT, where a narrow column cannot sustain propagation in cases where lateral support from the surrounding snowpack on a three-dimensional slope can. Additionally in the ECT, edges on the shovel blade often dig into a shallow weak layer cutting off the rest of the layer from the section under the shovel and preventing potential propagation. Simenhois and Birkeland (2006) also acknowledged this limitation. In most of these thin-slab cases, fractures that began in the PST arrested at a slab fracture, and often initiated in the ECT without subsequent propagation. Although these were the source of many false-stable predictions in the dataset, such shallow, soft slabs are rarely the cause of harmful avalanches in the Columbia Mountains (e.g. Jamieson and Johnston, 1993b).

PST and ECT results plotted against weak layer depth indicated that during the early stages of burial for the tested persistent weak layers, propagation propensity generally increased as the overlying slabs thickened. This was evident by the transition from PST SF and ARR results and ECTN results to more PST END and ECTP results. High propagation propensity was capable of persisting for much of the winter, while only fracture initiation became more difficult. Since very few wind-stiffened slabs were tested in the dataset, it can generally be assumed that due to increasing settlement and overburden load, slab density and stiffness increased with slab thickness (e.g. McClung and Schaerer, 2006). This is especially true in the layers immediately above and below the weak layer, as previously shown to favour propagation (e.g. B.C. Johnson, 2000; van Herwijnen and Jamieson, 2007b; Section 2.5).

It is the specific combined slab and weak layer characteristics with high propagation propensity that are of most interest to forecasters and recreationists. The studies discussed in Section 2.5 showed that thicker, harder slabs often favoured propagation, and that thin, soft, large-grained weak layers (typical of PWLs) were also conducive to propagation. Although weak layer variables were not explored in detail in this study, similar results relating slabs – and the slab and weak layer characteristics in combination –

to PST and ECT predictions of propagation propensity were found in the results presented here.

Since all of the data come from the Columbia Mountains, it is important to note that strengths and limitations determined from this dataset may predominantly apply to the soft snow at and below tree line in the Columbia Mountains and to surface hoar and crust-facet combinations commonly tested. For example, Simenhois and Birkeland (2009) found the ECT generally provided reliable results in snow up to 100 cm deep in the climates of the Colorado Rockies and New Zealand, where wind-stiffened slabs were commonly tested.

4.4.7 Summary

Both the PST and ECT have been shown to accurately predict propagation propensity in most cases within an expected range of skier-triggerable slab depths (30-70 cm), but showed particular strengths or limitations at un-validated sites outside that range. In shallow layers less than 30 cm deep, neither the PST nor ECT propagated many fractures, which coincided with the false-stable predictions of the PST in the validation study. Beyond that range only the PST was capable of testing propagation propensity in deep slabs. Validation of these deep PST results is attempted in Section 4.5. By learning the conditions in which a particular test is more appropriate, practitioners and recreationists can make more efficient and accurate snowpack observations.

4.5 Validation of Deep-Slab PST Results

Predicting the occurrence of deep-slab avalanches demands the evaluation of propagation propensity in deeply buried weak layers, for which the PST appears suitable based on commonly observed *propL* results in slabs between 70 cm and 250 cm thick (Section 4.4.2). Validating the PST results in deep slabs presented a unique challenge since skier-testing becomes dangerous and impractical at such depths, and naturally and accidentally triggered deep-slab avalanches are rare. For the following analysis slabs 70 cm or thicker are considered deep slabs. Seven such slabs exist in the slope-scale validation dataset – all between 70 and 90 cm thick – with mixed prediction accuracy. Aside from these limited results at validated sites, the interpreted results (*propL* or *propUL)* of deep-slab PSTs were compared to forecaster's expert ratings of propagation propensity to determine if a correlation existed. If so, further correlation could be sought between propagation lengths observed in the PST, and the expected extent of propagation (relative to start zone size) rated by the forecaster. The large area of Glacier National Park over which the ratings were considered to apply created a spatial scale issue when compared to test results from single sites within the park, and thus any correlations between tests and ratings were not expected to be high.

4.5.1 Data

The lead forecaster at Rogers Pass rated propagation propensity as *likely (*L), *equally likely* (E) or *unlikely* (U) almost daily for each of the persistent weak layers commonly tested throughout the 2008 and 2009 winters. Ratings were considered to apply to the average conditions on that layer throughout Glacier National park. Ratings began with initial burial and continued until the end of each season for the following layers: the 5 December (2007) crust and 26 January, 23 February and 9 March surface hoar layers in 2008 (Section 3.2.1); and the 25 December facets and 27 January, 22 February and 1 March surface hoar layers in 2009 (Section 3.2.2). Expert ratings were considered to apply to one day on either side of the rating date provided that new snowfall or drastic temperature change was not a factor between the two days. This generated 133 expert ratings of propagation propensity for a specific layer on a given day in which at least one PST from the same or adjacent day could be compared. In terms of deep slabs, 42 of these ratings were for layers that were buried deeper than 70 cm at the test site. Since most comparable days had two PSTs and sometimes more, 309 test results had an associated rating, of which 86 were in deep slabs (> 70 cm). No tests were performed on the 9 March SH within one day either side of an expert rating. Data supporting the comparisons are shown in Appendix C.

4.5.2 PST results compared to expert ratings of propagation

To determine if expert ratings could validate deep-slab PST results (> 70 cm) different comparisons were made that sought potential correlations. Some used all three ratings or only the L and U ratings, and compared them to individual test results or the *combined interpretation* of all test results on the same layer in one snowpit. The various comparisons are as follows:

I. L, E and U ratings versus the combined interpretation of PSTs;
II. L, E and U ratings versus individual PSTs;
III. Only L and U ratings versus individual PSTs, excluding any E ratings;
IV. L, E and U ratings versus combined interpretations or individual PSTs at Fidelity;
V. L, E and U ratings versus combined interpretations or individual PSTs on each PWL;
VI. L, E and U ratings versus combined interpretations of PSTs for layers shallower than 70 cm, between 70 cm and 100 cm, or deeper than 100 cm separately.

Likely or *unlikely* ratings were compared to the results of *propL* or *propUL* interpreted from individual PSTs and were marked *true* or *false* based on whether they matched. For combined interpretation of PST results, all results on the same layer within a pit had to agree on *propL* or *propUL* in order to match the respective L or U rating. If results were conflicting within a pit, they were interpreted as indicating propagation was *equally likely*. The E rating proved ambiguous and required some subjective interpretation when comparing to PST results. For instance, two conflicting PST results within the same pit could not be matched with L or U in combination, but could arguably be matched with E, as could results that required near-half cuts and propagated to the end. This allowed results with 45% cut, for example, to be matched with either L or E ratings which potentially introduced bias towards creating matches in the analysis. Other such comparative criteria that may have created biased matches were considered given that experience with the PST allows for a more subjective interpretation of certain results than the objective ’50%-cut rule’. The data were assessed with and without potentially biased interpretation.

The criteria used to determine matches between individual or combined PST results and expert ratings of propagation are listed below, and the results are given in Table 4.16:

- L ratings must have two *propL* PST results for a combined match
- U ratings must have two *propUL* PST results for a combined match
- E ratings can match conflicting PST results, PSTs in the 40%- to 60%-cut range propagating to END, or less than 45% cut with long propagation to ARR.
- Conflicting PSTs only match an E rating
- 40% -to 50%-cut to END PSTs can either match L or E ratings
- 50%- to 60%-cut to END PSTs can either match U or E ratings
- Less than 45% cut with long propagation to ARR can either match U or E ratings.

The only criteria that were unique to the comparison involving combined test interpretations was for cases of conflicting PST results, where propagation propensity from test results was interpreted to be equally likely. When individual tests were interpreted separately these conflictions were no longer a factor. This resulted in minimal differences between comparisons involving combined test interpretations and individual test interpretations in most cases, although both are given where available in Table 4.16. For example, combined interpretations of multiple PSTs on the 26 January surface hoar matched the expert rating on seven days and mismatched on two. Individually, 15 tests matched the rating while 12 were mismatches.

Table 4.16: Various methods of comparing expert ratings to the interpreted results of combined PSTs or individual PSTs on a specific layer within a given pit were used to determine the percentage of times ratings matched test results in different scenarios. Only rated/tested layers deeper than 70 cm are used except where noted.

RATINGS/TESTS COMPARED	**Combined PST interpretation**		**Individual PST interpretation**	
⇩ **MATCH:**	**TRUE**	**FALSE**	**TRUE**	**FALSE**
L, U, and E ratings:	**22 (52%)**	**20 (48%)**	**44 (51%)**	**42 (49%)**
Only L and U ratings:	-	-	**27 (55%)**	**22 (45%)**
Tests at Fidelity only:	**11 (69%)**	**5 (31%)**	**20 (77%)**	**6 (23%)**
By persistent weak layer:				
• *5 Dec. CR/FC*	*5 (71%)*	*2 (29%)*	*7 (78%)*	*2 (22%)*
• *26 Jan. SH*	*7 (78%)*	*2 (22%)*	*15 (56%)*	*12 (44%)*
• *23 Feb. SH*	*4 (67%)*	*2 (33%)*	*8 (80%)*	*2 (20%)*
• *9 Mar. SH*	-	-	-	-
• **All 2008 layers**:	**16 (73%)**	**6 (27%)**	**30 (65%)**	***13 (35%)***
• *25 Dec. FC*	*2 (33%)*	*4 (66%)*	*5 (38%)*	*8 (62%)*
• *27 Jan. SH*	*3 (37%)*	*5 (63%)*	*7 (47%)*	*8 (53%)*
• *22 Feb. SH*	*1 (25%)*	*3 (75%)*	*2 (25%)*	*6 (75%)*
• *1 Mar. SH*	*0 (0%)*	*2 (100%)*	*0 (0%)*	*4 (4%)*
• **All 2009 layers**:	**6 (30%)**	**14 (70%)**	**14 (35%)**	**26 (65%)**
By layer depth				
• Under 70 cm	16 (31%)	36 (69%)	-	-
• 70 to 100 cm	13 (65%)	7 (35%)	-	-
• Over 100 cm	9 (41%)	13 (59%)	-	-

SH = surface hoar; CR/FC = crust and facet combination; FC = facets

4.5.3 Discussion of deep-slab validation results

The objective comparison using only L/U ratings and *propL/propUL* results had a higher percent of matches than the comparison that included subjective PST interpretations to match E ratings, thereby discounting the correlation that may have arisen through biased interpretations. Test results at Fidelity show a higher percentage of matches than comparisons throughout all of Glacier National Park. This can be expected since Fidelity is the main source of snowpack and weather observations for the forecasters that rate propagation propensity and thus test results there will likely correlate better with ratings. However, the success at Fidelity may give some validation to the results in deep-slab PSTs

A limited number of comparisons existed for individual PWLs thus any apparent success on one particular layer is not likely significant, although some do show better results over others. This may have partially resulted from some layers having lower spatial variability across the region or showing more consistent propagation trends throughout the season making comparisons more likely to match. The results of 2008 are considerably better than the results of 2009. Since the test method and interpretation of results were the same in both seasons, it is suggested that the accuracy of the ratings are highly dependent on the general 'ease' or confidence of forecasting propagation on those layers, which is influenced by spatial and temporal variability exhibited by the layers, and the accumulation of those differences between seasons.

When categorized by depth, layers 70 to 100 cm deep had the best match rate while layers shallower than 70 cm had the worst. This may be explained by a stronger dependence of propagation potential on depth in the early stages of layer burial as described in Section 4.4.6, after which it may show more consistency. Layers deeper than 100 cm showed a decreased match rate compare to those between 70 and 100 cm. In layers at these depths, observations by the forecaster are rarer thus making rating propagation more difficult.

4.5.4 Summary

Comparisons of PST results to expert ratings of propagation propensity using the full dataset showed the two variables matched on only about 50% of days, even with the considerable bias introduced in the interpretation of *E* ratings. Tests performed at Fidelity matched expert ratings on an improved 69% and 77% of days depending on whether PSTs were interpreted together or individually, suggesting that Fidelity study sites may be representative of regional conditions. Comparisons on individual layers ranged from 0% matching to 80%, although too few comparisons exist to draw any kind of significant conclusions as to why they differed. Layers buried between 70 cm and 100 cm deep showed more matches than shallower or deeper ranges, although limited data supports this.

The original intent of this analysis was to assume that the expert ratings were *always correct*, and therefore the PST could be validated against them. Given the generally poor correlations between all ratings and all test results revealed in Table 4.16, validating deep-slab PSTs using expert ratings proved to be rather unsuccessful, indicating perhaps that snowpack conditions throughout Glacier National Park proved too variable to be captured in a single rating or in a single test.

4.6 Regional Validation of PST Results

In this section, local PST results at representative study sites are compared to avalanche activity reported in the InfoEx over the 2008 and 2009 winters in the Columbia Mountains using the methods described in Section 3.8. The present analysis depends on two inseparable hypotheses: that the selected study sites at Mt. Fidelity and Mt. Abbott in Glacier National Park, and Mt. Saint Anne near Blue River, are representative of regional weather and snowpack conditions; and that the PSTs performed at those sites are indicative of regional trends in propagation propensity.

Numerous avalanches reported in the InfoEx started in the alpine where wind and sun exposure can alter snowpack characteristics, although most started in an elevation band comparable to test sites. Test sites were near tree line and ranged between 1600 m and 2100 m in elevation, while the majority of avalanche start zones ranged between 1700 m and 2600 m in elevation. The study sites covered all aspects and varied in slope inclination from 0^{o} to 40^{o}, although most slopes were in the range of 25^{o} to 35^{o}.

Table 4.17 shows the eight PWLs of 2008 and 2009 that were tested regularly at study sites with the PST, and which produced significant avalanche activity throughout the Columbia Mountains. Each layer was reported regularly in the InfoEx and showed persistent avalanche activity for more than ten days, which fits the definition of a regional PWL described by Hägeli and McClung (2007). The active period of the layer began with the first potentially harmful avalanche (1.5+) and ended when no more avalanches were reported in the InfoEx. Since many operations stopped reporting in late March or early April (e.g. Hägeli and McClung, 2007), the active period may not be accurate but allows for relative comparisons between PWLs and with peak AAI. Peak AAI draws attention to the scale of propagation propensity associated with each layer. For example the 23 February 2008 surface hoar produced numerous large avalanches over a long period of time, generating a consistently high AAI and the largest single day peak (4,336) of all eight evaluated layers.

Table 4.17: Eight significant PWLs from the winters of 2008 and 2009 were tested locally with the PST and produced significant and prolonged avalanche activity throughout the Columbia Mountains. The PWL burial date and predominant grain type are given, along with the period over which avalanche activity was reported and the peak avalanche activity index (AAI) on each layer. The number of PSTs performed at regionally representative sites is also listed, along with the percentage of days within the active period on which a PST was performed

PWL date ID	PWL type	Active Period	Peak AAI	PSTs	Percent of days in active period tested with the PST
5 Dec. 2007	CR/FC	60 days	510	23	15%
26 Jan. 2008	SH	59 days	1,114	19	12%
23 Feb. 2008	SH	46 days	4,336	34	26%
9 Mar. 2008	SH	20 days	190	18	30%
25 Dec. 2008	FC	85 days	1,542	13	7%
27 Jan. 2009	SH	53 days	595	26	25%
22 Feb. 2009	SH	29 days	1,035	20	31%
1 Mar. 2009	SH	19 days	548	15	32%

The 9 March, 2008 and 1 March, 2009 SH layers were omitted from further analysis due to the limited time over which the layers were tested and avalanche activity recorded before the end of the field season.

4.6.1 PST results and regional avalanche activity over time

Some peaks in the AAI occurred during storm cycles when the buried PWL was subjected to rapidly increasing load, producing many small and some large avalanches. Alternatively, some peaks were the result of one or more large avalanches triggered artificially, particularly later in the buried life of the layer. Since peaks in AAI were heavily influenced by triggering ease and storm cycles they are not an exclusive indicator of fluctuating propagation propensity. Propagation propensity *may* increase rapidly if the overlying slab stiffens during a storm cycle, but typically grows and decays over a longer period of time (e.g. van Herwijnen and Jamieson, 2007b; refer to Section 1.5). *However*, propagation propensity *must* be high for large peaks in the AAI to occur, since propagation is required for slab-avalanching. In other words, an increase in the AAI must be accompanied by high propagation propensity but not vice-versa; and a decrease in the AAI may be due to reduced triggerability. This means the peaks in AAI only offer 'glimpses' into the propagation propensity of a layer at specific points in time when a trigger is present. Propagation propensity likely remains high between successive peaks in activity, and may arguably be 'growing' in the time leading up to the first peak and possibly waning in the time after the last observed activity. This process is idealized in Figure 4.10.

When evaluating PST results on the regional scale, success was attributed to tests that indicated *propL* during cycles of high avalanche activity. PSTs that indicated *propL* when avalanche activity was *low* were not necessarily incorrect since avalanche activity is also a function of triggerability. In other words, buried PWLs that are generally out of skier-triggerable range and are not being subjected to new snowfall loads will not produce many avalanches, even though the propagation propensity in that layer may at the time be high. For the same reason, PST results that indicated *propUL* during low avalanche activity *appear c*orrect – and may be, but it must also be considered that regional propagation propensity could have been high and triggerability low at the time. PSTs that indicated *propUL* during high avalanche activity were the only conclusively incorrect regional evaluations of propagation propensity, although spatial variability allows any incorrect (or correct) PST to be the result of locally different conditions.

Since propagation propensity often grows and decays slowly (e.g. van Herwijnen and Jamieson, 2007b), PST results may be indicative of conditions within a number of days either side of the peaks in avalanche activity.

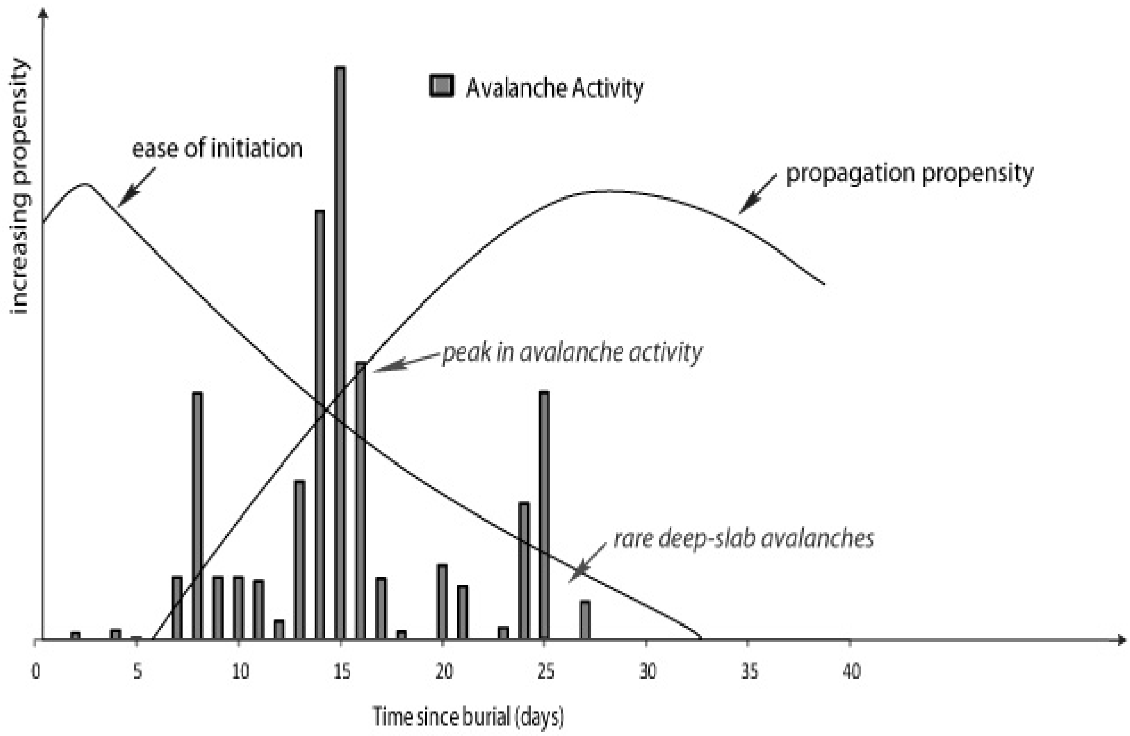

Figure 4.10: Idealized representation of fictitious daily avalanche activity over the active period of a PWL and the growth and decay of initiation (triggering) ease and propagation propensity over the same period (see Figure 1.5). The peak in daily avalanche activity may be the result of many smaller avalanches a few large avalanches, whereas the later peak is more likely the result of one or a few large destructive avalanches with wide propagation.

Figures 4.11 to 4.16 show the regional avalanche activity and PSTs performed at representative sites for the first three significant PWLs of each of the 2008 and 2009 winters. The March PWLs from each season are not shown. Results are plotted against time-since-burial and thus show trends in propagation propensity both in terms of regional avalanche activity and in terms of PST predictions. The active period of a layer often extended with intermittent activity for many days beyond the last PST result on the layer; however the AAI for each layer is not plotted beyond two weeks from the last PST result.

4.6.1.1 5 December 2007 crust

PST results and avalanche activity on the 5 December crust are plotted over 120 days-since-burial in Figure 4.11. The avalanche activity index peaked 17 days after burial at 510, with three subsequent and smaller peaks around 33, 43, and 69 days after burial. PST results on days surrounding the first, third, and fourth peak mostly agreed with the high propagation propensity reported on that layer during those times. However, leading up to the second peak, four of six PSTs indicated *propUL.* Three of five PSTs after the fourth peak in activity indicate *propUL* at a time when no avalanche activity was reported on the layer. Although these results may be indicative of a decline in propagation propensity, this cannot be stated with certainty since only triggerability may have declined.

4.6.1.2 26 January 2008 surface hoar

PST results and avalanche activity on the 26 January surface hoar layer are plotted over 30 days in Figure 4.12. The five-day running average AAI demonstrated a continuous growth and decay over 25 days despite daily fluctuations in activity, with a single-day peak just over 1,100 on day 15. During the first ten days after burial, PST results mostly predicted *propUL*, which slowly transitioned to *propL* results by the peak in AAI between the 15th and 18th day (note the shorter cut to ARR and slightly shorter cut to END *propUL* results on the 11th day representing a possible transition towards *propL*).

4.6.1.3 23 February 2008 surface hoar

PST results and avalanche activity on the 23 February surface hoar are plotted over 40 days in Figure 4.13. This layer showed the highest single-day peak in activity (4,336) and significant avalanche activity was reported almost daily through to the end of March. Even during the lower-activity period between the two peaks the average AAI hovered around 250, which represented considerable avalanche activity – perhaps numerous size twos and a size three daily. PST results during that time were mixed, although 15 of 18 PSTs within three days either side of the two peaks predicted *propL*.

4.6.1.4 25 December 2008 facets

PST results and avalanche activity on the 25 December facets are plotted over 40 days in Figure 4.14. Regional avalanche activity on the layer peaked (1,542) at day 15 during a prolonged storm cycle in early January. Of the nine PSTs performed during that event, seven predicted *propL.* After day 30, PST results were split, perhaps because triggering became difficult. Nevertheless, the validity of either *propL* or *propUL* cannot be inferred.

4.6.1.5 27 January 2009 surface hoar

PST results and avalanche activity on the 27 January surface hoar are plotted over 75 days in Figure 4.15. Activity on this layer peaked early, but the AAI never exceeded 600. Avalanche activity remained low on this layer with the exception of another peak 38 days after burial. Most PSTs performed on this layer were on the days between the two peaks and primarily predicted *propL*. Six of seven PSTs performed on the days immediately following the first peak in activity predicted *propL*, although the *propUL* was predicted on the day of and the day before the peak.

4.6.1.6 22 February 2009 surface hoar

PST results and avalanche activity on the 22 February surface hoar are plotted over 45 days in Figures 4.16. Similar to the 27 January surface hoar, avalanche activity peaked early and was not well predicted by the PST. Four *propL* results around the tenth day coincided with some significant avalanche activity around the same time. Although PSTs performed during the periods of lowest avalanche activity indicate *propUL*, the validity of these results cannot be assumed since natural or artificial triggers may not have existed at the time.

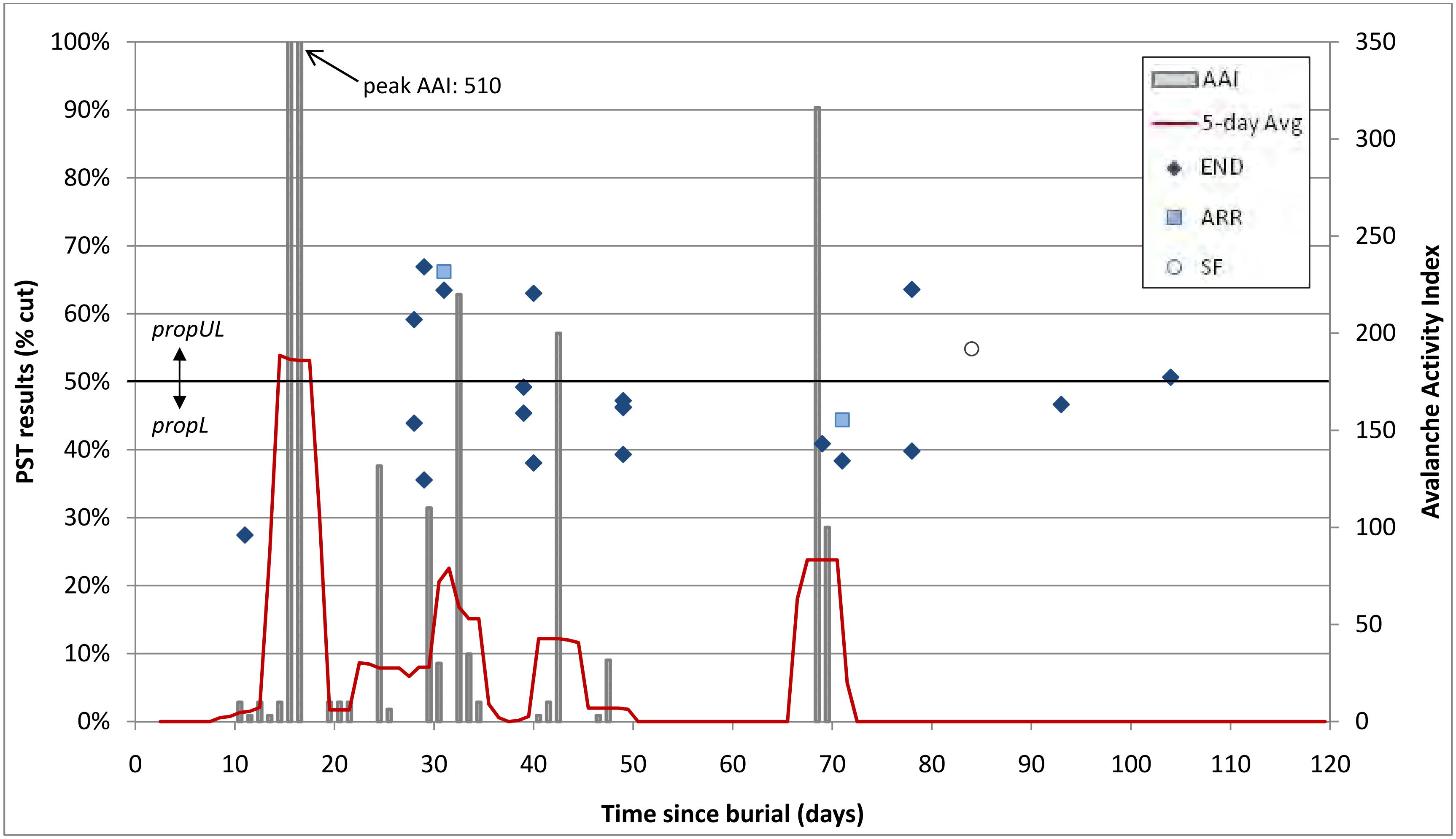

Figure 4.11: PST results (END/ARR/SF) and regional avalanche activity (AAI) over time on the ***5 December 2007 crust****. A centred five-day moving average AAI is also plotted. Four main peaks in activity are apparent, although avalanche activity was generally low. PST results were predictive of the first, third and fourth peak in activity.*

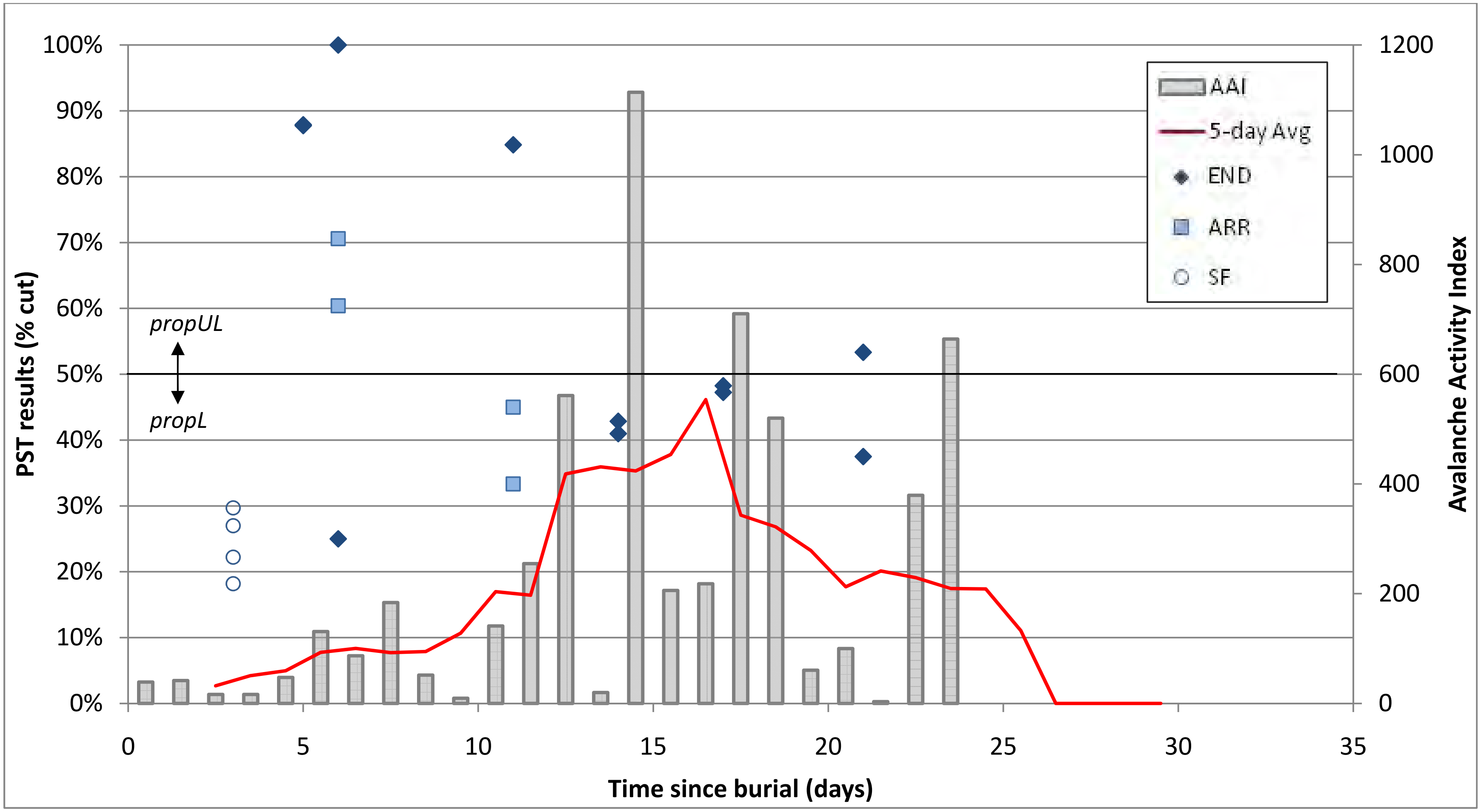

Figure 4.12: PST results and regional avalanche activity over time on the ***26 January 2008 surface hoar layer****. Avalanche activity fluctuated from day-to-day but showed an average growth and decay over approximately 25 days. PSTs mostly predicted propUL when low avalanche activity was recorded in the first ten days after burial. As avalanche activity peaked around 15 – 18 days after burial, PST results were indicating propL.*

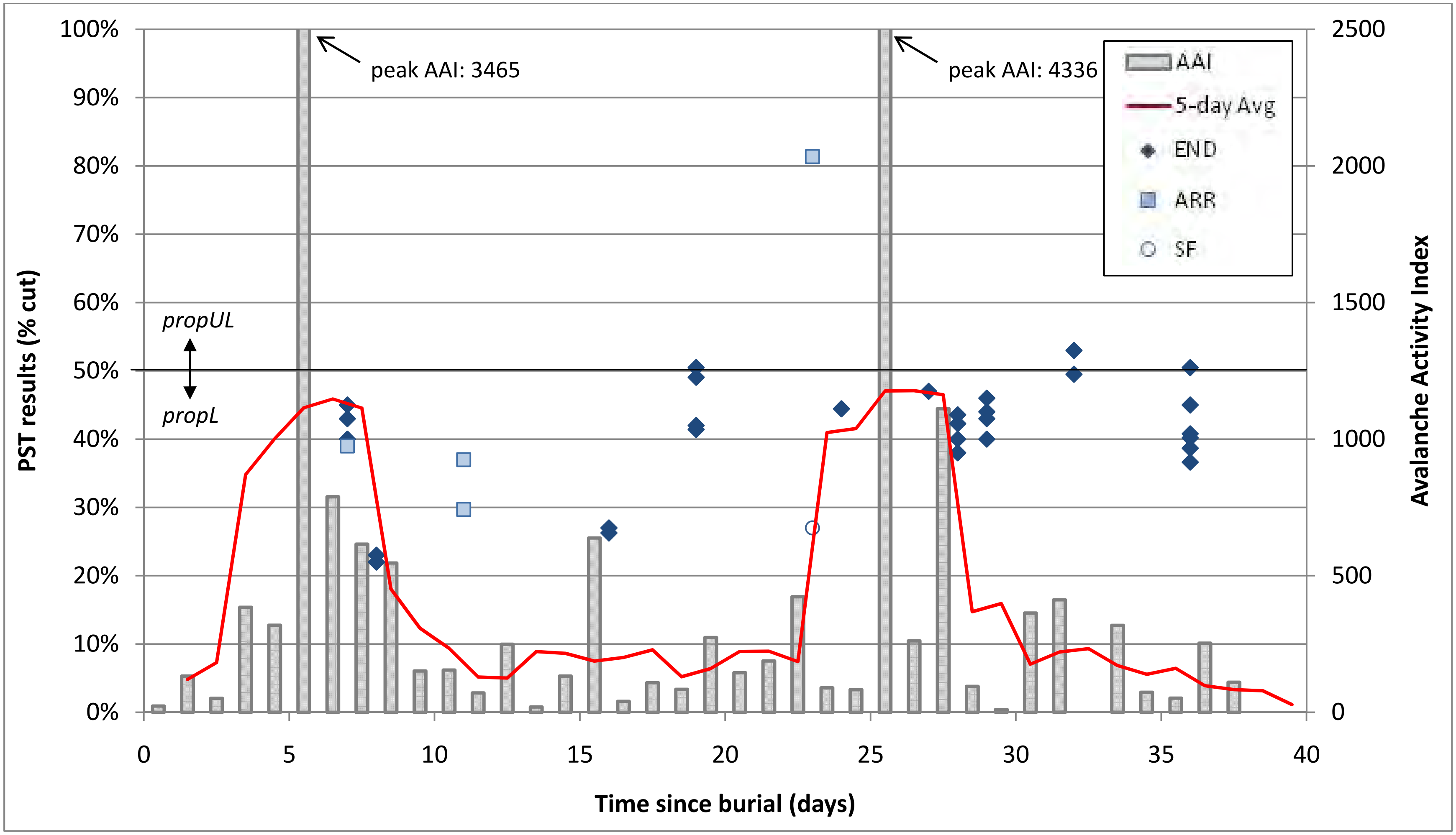

Figure 4.13: PST results and regional avalanche activity over time on the ***23 February 2008 surface hoar layer****. Two major peaks in activity were recorded during which PST results were predicting propL. In the period of fluctuating activity between peaks, PST results were mixed perhaps representing the variability over that period.*

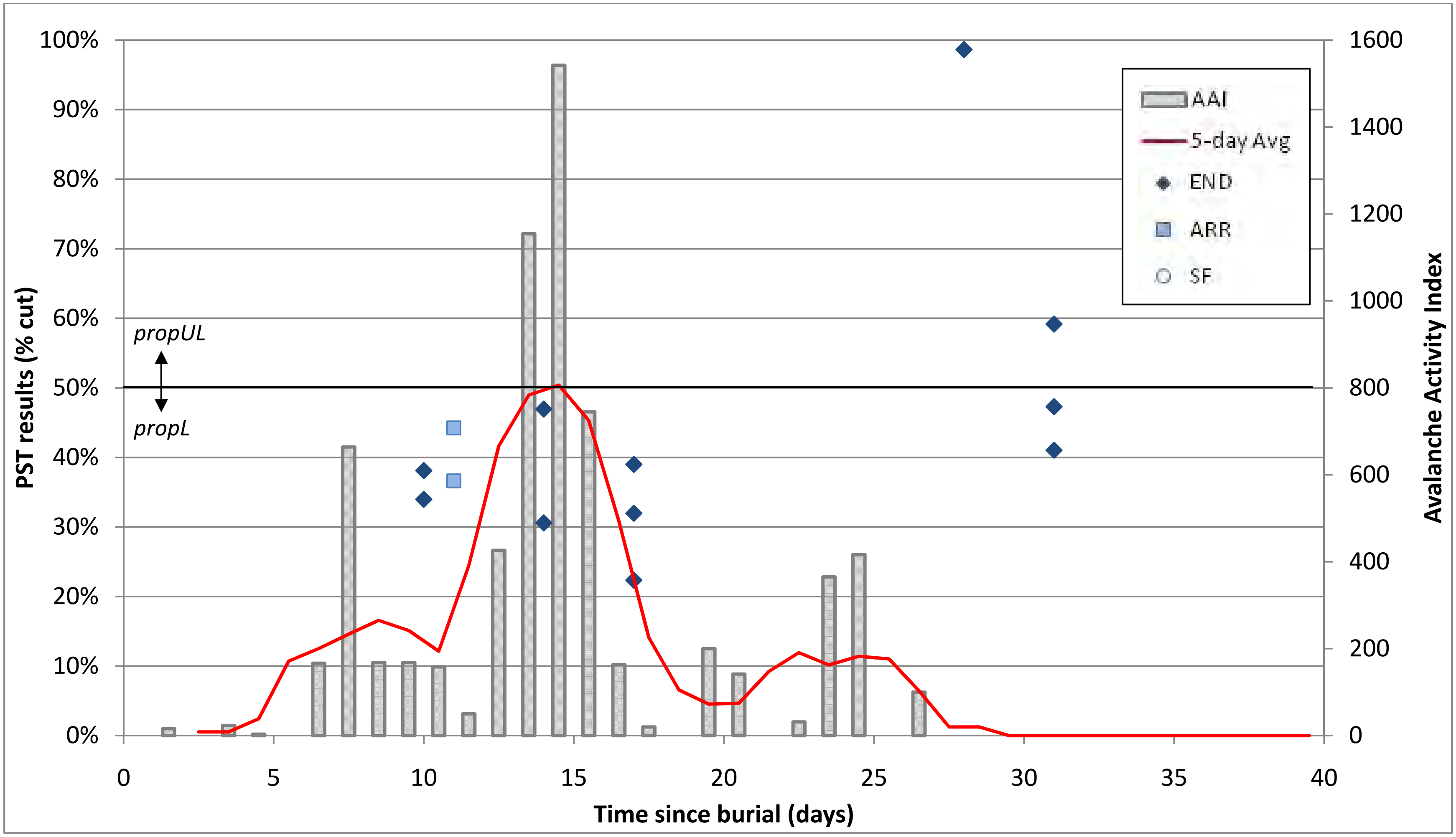

Figure 4.14: PST results and regional avalanche activity over time on the ***25 December 2008 facet layer****. Avalanche activity peaked 15 days after the layer was buried, but intermittent activity was reported through to the end of March. Seven of nine PSTs performed within a few days either side of the peak indicated propL.*

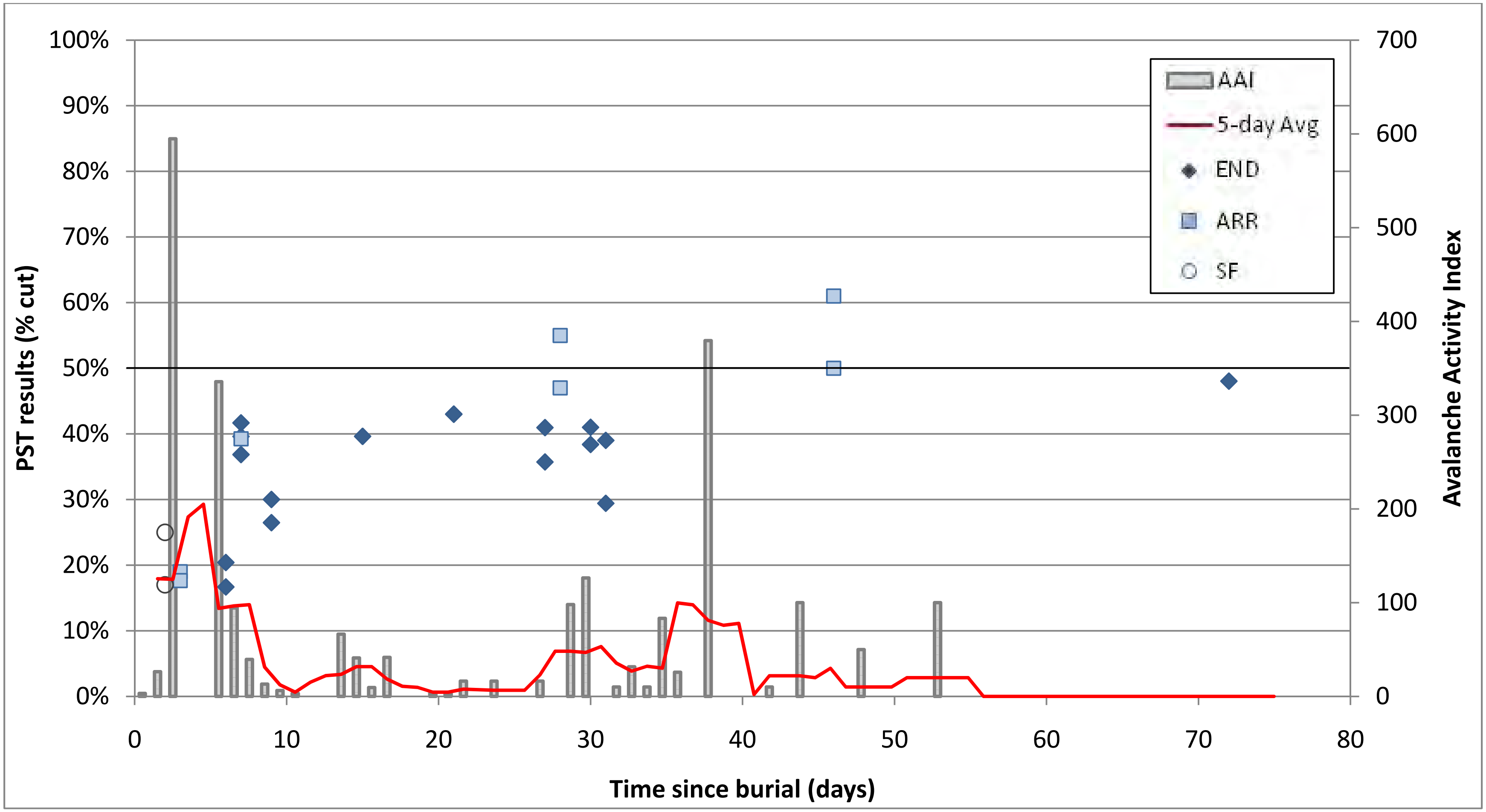

Figure 4.15: PST results and regional avalanche activity over time on the ***27 January 2009 surface hoar layer****. Regional avalanche activity peaked early on this layer and was not well-predicted by the same- and previous-day's PSTs, although six of seven PSTs in the days shortly following the peak indicated propL. The small spike in activity around day 30 is well-predicted (6 of 8) and no tests were performed within a few days either side of the peak on day 38.*

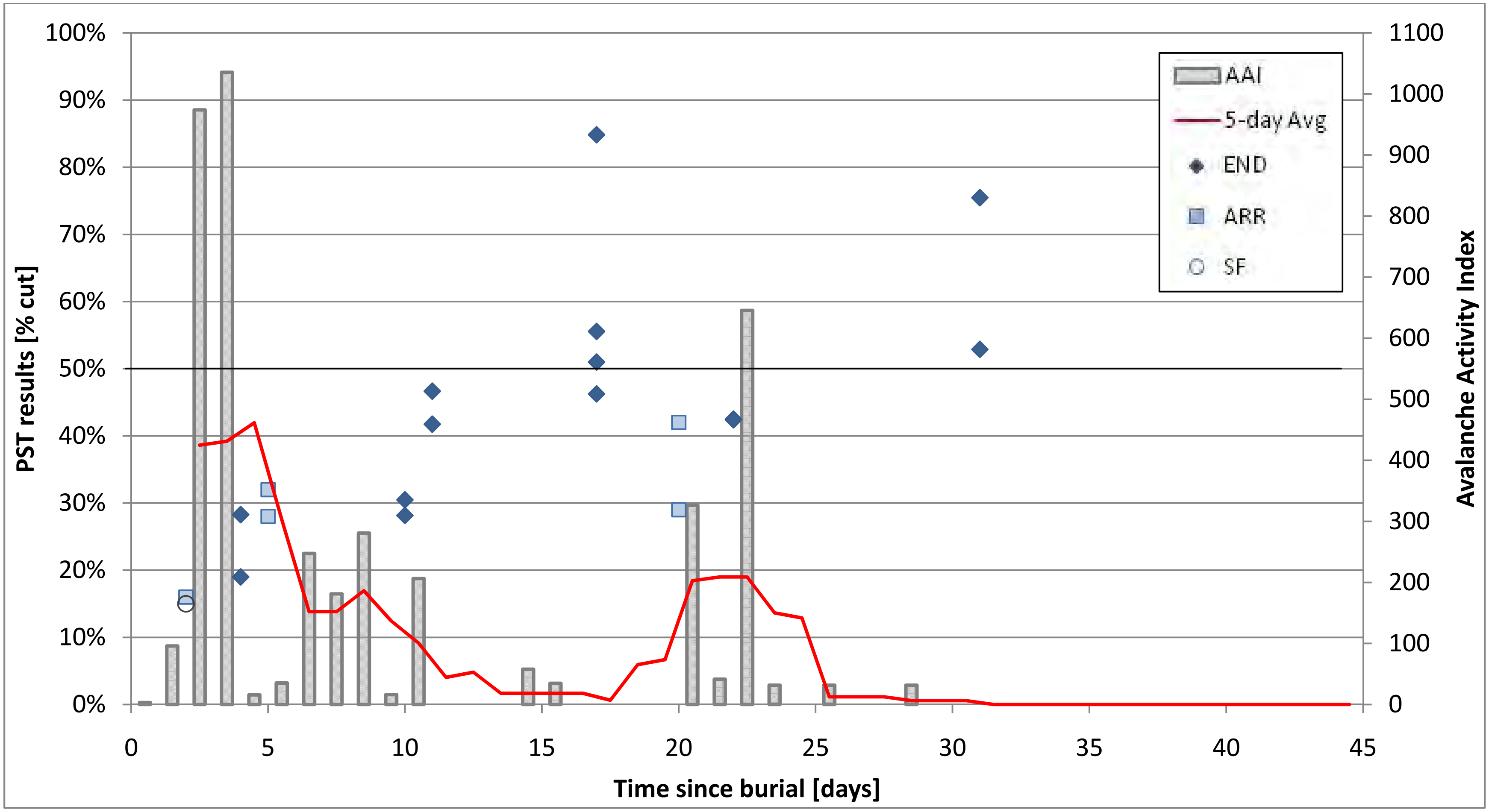

Figure 4.16: PST results and regional avalanche activity over time on the ***22 February 2009 surface hoar layer****. Activity on this layer also peaked early, with a successive peak around day 23. Only two for six PSTs performed during the peak activity predicted propL and only one of three predicted propL during the later peak. A few PSTs during the lowest avalanche activity around day 18 and beyond day 30 indicated propUL.*

4.6.2 Discussion of regional validation results

Because of the spatial variability within weak layers, within properties of the overlying slab, and in their combined development over time, studies of avalanche phenomena on the regional scale often arrive at general or limited conclusions (e.g. Schweizer et al., 2003b; Jamieson et al., 2007). The objective of the current study was to evaluate the predictive merit of the PST on the regional scale despite the limited success of such previous studies on that scale. This was based on the widespread prevalence throughout the Columbia Mountains of four significant PWLs in each of the winters of 2008 and 2009 (see Section 4.6.1), and on the hypothesis that propagation propensity may show less spatial variability than other snowpack characteristics (e.g. Johnson and Birkeland, 2002; Campbell and Jamieson, 2007; van Herwijnen et al., 2009).

The present analysis had several limitations:

I. The AAI was expected to be a strong indicator of when propagation propensity was high, but not necessarily low; and may be more indicative of triggering on some days (particularly during large storm cycles);
II. Only *propL and propUL* results surrounding peaks in the AAI could be validated; low AAI alone could not validate either *propL* or *propUL* results.
III. The PST performed at study sites were expected to be indicative of propagation propensity trends on the regional scale, which necessitated the inseparable and unproven hypothesis that the selected study sites are representative of snowpack conditions on the regional scale.

The PST often indicated high propagation propensity during periods of increased or peak regional avalanche activity for most of the PWLs that experienced significant avalanche activity throughout the Columbia Mountains. Despite the limitations of this study, the success is nonetheless encouraging. Even though the *propUL* results during time of low activity could not be validated, the fact that such occurrences exist brings up another valuable observation: that if the PST only predicted *propL* throughout the life of the layer it could arguably not be proven wrong, but since *propUL* results are inter-dispersed throughout, it shows the PST is not ``hedging`` results. To evaluate *propUL* PST results, storm events could be plotted throughout the season along with avalanche activity so that any large storm event that *did not* produce large peaks in avalanche activity on a layer could arguably validate *propUL* results. This was beyond the scope of this study and is recommended for future analysis.

The two PWLs on which regional avalanche activity peaked early after burial (27 January and 22 February, 2009) were poorly evaluated by the PST. This likely coincides with evidence presented in the slope-scale validation (Section 4.2) that the PST performs poorly in shallow, soft slabs typical of early avalanche cycles on a layer.

4.6.3 Summary

PST results from local study sites and regional avalanche activity reported in the InfoEx were compared over time-since-burial on three PWLs in each of 2008 and 2009. Despite the limitations of evaluating regional conditions with local tests, the PST appeared to successfully indicate propagation propensity was high during periods of elevated avalanche activity on most of the PWLs observed throughout the Columbia Mountains each winter. This appears to support the hypothesis that trends in propagation propensity on significant, region-wide PWLs may show less spatial variability at a low resolution than other snowpack characteristics, and that the PST can be indicative of these trends.

5. CONCLUSIONS

5.1 Thesis Conclusions

The recent experimental and theoretical devotion of avalanche researchers and practitioners to the process of fracture propagation in buried weak layers has led to many new and exciting developments, of which the PST is one. The work of Gauthier (2007) and Gauthier and Jamieson (2006a, 2007a, 2008a) established the standard PST geometry and method and contributed some initial slope-scale validation results used here. This study expanded on their slope-scale validation by more than tripling the dataset, and presented new and distinct experimentation with scaled column lengths, field-based comparisons with the ECT, and original efforts to validate the PST for deep slabs using forecaster's expert ratings, and on the regional scale using avalanche activity data. In the winters of 2008 and 2009, 783 standard PSTs, 183 Extended Column Tests (ECT), and 231 experimental short-scaled PSTs were performed for this study, including 99 PSTs at 53 validated sites. Almost all testing was in the Columbia Mountains, with a transitional snow climate and primarily crust-facets and surface hoar PWLs, as described in Hägeli and McClung (2003).

Separated by the objectives outlined in Section 1.10, the conclusions of this thesis are as follows:

5.1.1 Slope-scale validation of the standard PST

The results of 169 PSTs on 84 site-layers at 75 validation sites of confirmed propagation or confirmed initiation without propagation from 2007-2009 were combined in Section 4.2 with the following results:

- 75% of all validated PSTs made correct predictions about propagation propensity, while 22% falsely predicted low or no propagation propensity and 3% falsely predicted high propagation propensity. Results were comparable from year to year, with overall accuracy ranging from 61% to 82% and false-stables ranging from 14% to 37%. False-unstable results were consistently low between 0 and 5%;

- The calculated True Skill Statistic and Unweighted Average Accuracy of the PST for the current dataset were 0.58 and 0.79 respectively, which compared to – and often exceeded – the predictive skill of other standard snowpack tests and assessment methods (calculated from other sources), especially when only propagation propensity was tested and not initiation ease.
- False-stable results of the PST were higher than for other tests, and commonly occurred in shallow, soft slabs or at day-old avalanche test sites. In shallow, soft slabs, slab fracture (SF) results were commonly observed in the test where the natural, 3D slope appeared more able to sustain propagation before the slab fractured. When day-old avalanche sites were omitted, overall accuracy, TSS and UAA improved to 83%, 0.69. and 0.85 respectively suggesting that in some cases the flanks of day-old avalanche sites may no longer – or never did – have the high propagation propensity that existed when the slope avalanched.
- PST results proved to be highly reproducible at validation sites, with only 15% of sites having both correct and incorrect test predictions.
- The small validated dataset of ECT results collected in 2008 (*n* = 50) showed an overall accuracy of 58%, TSS of 0.23 and a UAA of 0.62, which is much lower than other published results (e.g. Simenhois and Birkeland, 2009).

5.1.2 Scaling the PST below 1 m with weak layer depth

Subjective field analysis of 231 short-scaled PSTs and validation of 73 short-scaled PSTs at sites of propagation or initiation without propagation yielded the following results:

- Short-scaled PSTs were performed in a full range of slab thicknesses between 10 cm and 93 cm. In longer/deeper columns the test performed similarly to the standard PST method, but in weak layers beneath shallow, soft slabs the short-scale PST proved difficult to interpret, with many columns pivoting around the saw instead of propagating the fracture, and many more within +/- 2 cm of 50% cut.

- Validation results indicated low overall accuracy with numerous false predictions and results that could not be interpreted with confidence. Although the short-scaled PST gave correct predictions in shallow, soft slabs at two sites where the standard PST failed, it was generally problematic in such snowpack conditions.
- The standard PST method of 1 m long columns and scaling with weak layer depth *above* a metre is a much better predictor of propagation propensity and is recommend along with the recording standard presented in Section 3.3.

5.1.3 Practical results of side-by-side PST and ECT comparisons

Various comparisons of side-by-side PSTs and ECTs in the same snowpit were made and showed the following advantages and limitations of each test:

- The ECT could be performed without any prior knowledge of weak layers buried in the snowpack whereas the PST required the user to identify a weak layer of interest prior to performing the test. Both test methods were capable of producing multiple results in the same test column; the ECT by continued tapping, and the PST by cutting successive weak layers starting with the deepest.
- In both test methods, a minimum slab thickness of approximately 30 cm over a PWL was required to produce propagation *likely* results in the typical Columbia Mountain snowpack, with shallower tests commonly ending in PST_SF, or ECTN.
- The ECT appeared capable of indicating both initiation ease (triggerability) and propagation propensity in the same test reliably up to a depth of 70 cm and occasionally up to 95 cm, beyond which initiating a fracture in the deep weak layer via surface loading became rare. In the range of 30-70 cm, the ECT score depended on depth but the results of ECTN or ECTP did not show a dependence on depth.

- The PST was capable of predicting propagation propensity in deeply buried weak layers, with tested depths of up to 175 cm in 2009, up to 250 cm in 2008 and up to 285 cm in 2007 (Gauthier, 2007).
- In the PST, frequent END results occurred after 80-100% of the column had been cut, and less frequently between 60-80%, suggesting that the free upslope end of the column may attract the fracture after more than 80% of the column is cut.
- In direct comparisons at single test-sites, both test methods commonly agreed on propagation propensity when two of each were performed. In most cases where the PSTs and ECTs gave conflicting predictions the ECT resulted in ECTX whereas the PST predicted *propL*, primarily in deeply buried weak layers (> 70 cm).

5.1.4 Validation of deep-slab PST results with expert ratings

Validation of deep-slab PST results (> 70 cm) was attempted through comparison with forecasters' expert ratings of propagation propensity, with the following results:

- The rating of *equally likely (E)* proved ambiguous and required subjective interpretation when comparing to PST results, potentially introducing some bias towards generating matches.
- Using all sites and all layers deeper than 70 cm, PST results matched expert ratings approximately 50% of the time regardless of whether the *E* rating was included or not, or whether PSTs were interpreted individually or in combination from within the same snowpit. Matches in deep layers were more numerous than matches in layers shallower than 70 cm, perhaps indicating less spatial and temporal variability and thus easier rating and more consistent testing of propagation propensity as the weak layers were more deeply buried.
- When only tests from study sites at Mt. Fidelity were included, results matched ratings between 69% and 77% of the time, which was not unexpected since Fidelity is the main source of snowpack and weather information used by the same

forecasters rating propagation propensity. However, this result may help confirm that Fidelity is representative of conditions in the park, and that PST results at Fidelity are indicative of these average conditions.

- The number of matches varied substantially between individual layers, and more ratings and tests matched in 2008 than in 2009, perhaps indicating the spatial and temporal variability associated with different layers, in terms of both test results and the challenge of applying one rating to the average conditions. Too few data were available to make significant conclusions.

5.1.5 Regional validation of PST results

PST results at assumed regionally representative sites were compared to avalanche activity reported on six PWLs observed throughout large parts of the Columbia Mountains in 2008 and 2009, producing the following results:

- The significant PWLs of 2008 and 2009 used in the analysis showed consistency with the typical annual trends reported for PWLs in the Columbia Mountains by Hägeli and McClung (2003, 2007), giving some credence to regional validation results.
- PSTs were performed at the assumed representative sites on between 7% and 32% of days over which each PWL produced any reported avalanche activity, in most cases representing substantial temporal spacing and coverage within the active period.
- The PST often indicated high propagation propensity during periods of increased regional avalanche activity for most of the PWLs that experienced significant avalanche activity throughout the Columbia Mountains. The main exceptions were the two layers that showed peak avalanche activity early after burial, perhaps corresponding with false-stable predictions observed in shallow, soft slabs during slope-scale validation.
- Results were not expected to be strong given the large spatial variability issues over such a region and the mixed success of previous studies on regional validation (e.g.

Schweizer et al., 2003b; Jamieson et al., 2007). In addition, the limitations associated with assuming study sites were regionally representative, using avalanche activity as a proxy for regional propagation propensity, and being unable to validate test results during times of low activity were considered.

5.2 Application of the PST to Slab-Avalanche Forecasting

Reviews of avalanche accidents show that most fatal slab avalanches involve the initiation and propagation of fractures within persistent weak layers in the snowpack (e.g. Jamieson and Johnston, 1992), most commonly in the skier triggerable range of approximately 30 - 70 cm in depth, and occasionally under much thicker slabs. Because of this, initiation and propagation propensity of specific slab and PWL combinations are of great interest to practitioners and forecasters. Standard snowpack tests such as the CT or RB provide some insight into propagation propensity through observations of fracture character (e.g. van Herwijnen and Jamieson, 2002), shear quality (e.g. Johnson and Birkeland, 2002), or release type (e.g. Schweizer and Wiesinger, 2001), although they are all dependent on surface loading to initiate weak layer fractures and are thus likely better indicators of triggering ease than propagation propensity explicitly. Propagation propensity was defined for this study as the process following fracture initiation in which the fracture advances rapidly and independently of the initiation energy source and is instead driven by the energy sources inherent in the snowpack itself. The PST method explicitly tests the ability of a slab and weak layer to *sustain* fracture propagation independently of the initiation energy, and is thus a unique and validated addition to the forecasting toolbox (and for researchers e.g. van Herwijnen and others, 2008; Sigrist and Schweizer, 2007). The ECT also appears to be a capable and reliable predictor of propagation propensity, or at least the propensity for triggering and propagating a fracture together, although it is also depth limited. For example, when ECTX is recorded, it implies 'triggering is unlikely and propagation is unknown' rather than 'propagation is unlikely'.

The specificity of the PST method enables forecasters to target critical weak layers within the snowpack not only to gain insight on the immediate, slope-scale propagation propensity of the layer, but also to track the slab and layer's evolution over time. van Herwijnen and Jamieson (2007b) adequately described this process as the growth and decay of avalanche activity based on the relationship through time of initiation propensity and propagation propensity (Figure 1.5). Where other tests that depend on surface loading can be valuable in that they identify initiation propensity and perhaps propagation propensity in various layers in the upper snowpack, they can also distract observers from the most critical layer with potential for widespread propagation, and may fail to induce fracture in such layers once depths exceed 70 to 100 cm. With experience, practitioners and recreationists can use the PST to specifically test the propagation propensity of PWLs not only when surface-initiation propensity is high, but perhaps more importantly when it is not.

Simenhois and Birkeland's (2008a) observations of weak layer fractures propagating more readily from thin slabs to thick slabs within test columns agrees with observations of skiers triggering weak layers under thin slabs which then propagate under thicker surrounding areas to release large and often fatal avalanches. Jamieson and others (1998) also reported *natural* avalanches initiating from shallow spots and propagating under deep slabs to release large avalanches. These are important reasons why the PST is a valuable tool for assessing deep weak layers at snowpit sites where the potential exists for locally thin spots on the surrounding slopes. Although expert ratings of propagation propensity could not convincingly validate deep-slab PST results, it is reasonable to expect that the PST is equally as indicative of slope-scale propagation propensity in deep slabs as it is in the more commonly skier-triggerable range. Performing the PST at deep-slab avalanches is recommended as an improved means of validating the PST in deep slabs, although such data are difficult to obtain.

The PST has been shown to occasionally perform poorly in shallow, soft slabs. This appears to be a persistent limitation of the test method, although high propagation propensity in such conditions is infrequent and rarely consequential in terms of human risk. In these conditions, other snowpack tests such as the compression test and the associated fracture character (or shear quality) may be more appropriate for assessing stability. Furthermore, forecasting frequent soft-slab avalanches after a recent snowstorm is relatively well understood and straight forward compared to predicting or anticipating the potential for rare but deadly deep slab avalanches.

Performing regular slope-scale snowpack tests is often impractical for forecasters and thus large gaps in time and space can exist between point observations. For this reason, forecasting often involves extrapolating the results of snowpack tests beyond the slope scale and across time, which can be subject to challenging issues with spatial and temporal variability. The regional validation results of Section 4.6 gave encouraging evidence that trends in PST results may be suitable for such scaling across time and space, at least in terms of propagation trends in significant PWL's that have widespread regional prevalence for long periods of the winter. This was in agreement with observations by Jamieson and other (2007) that trends in stability indices rather than specific daily values correlated better with avalanche activity. Since regional forecasts such as CAC Public Avalanche Bulletins are heavily based on the existence of such regional PWLs, PSTs at well-selected representative sites may be an appropriate means of evaluating their (changing) potential for propagation. Although spatial variability could be high between one slope and the next or one drainage basin and the next (high resolution), large scale PWL formation and resulting trends in avalanche activity may produce a more homogeneous pattern of propagation propensity on a larger, lower-resolution regional scale as Hägeli and McClung (2003, 2007) have shown for other PWL characteristics. In fact, McClung (2000) argued that as the spatial scale decreases, forecasting becomes more difficult and the need for accuracy increases, which implies that at a large scale, identifying trends in snowpack behaviour may help improve regional forecasting. Hägeli and McClung (2004) added that the challenge of forecasting lies in evaluating the variable that appropriately represents a process at the scale of interest. Perhaps if propagation propensity on a wider regional scale is the relevant variable for PWLs, then the PST is an appropriate tool for measuring it.

Many field-based tools exist to aid forecasters in gaining a better understanding of the snowpack conditions pertinent to avalanche formation and release, particularly the snow-profile and snowpack (stability) tests. The PST is an additional tool with its own unique advantages and limitations. Understanding when to use the PST and when to try other tests can greatly increase the efficient gathering of snowpack information to improve forecasts.

5.3 Recommendations for Future Research

The following avenues of future research are recommended to address knowledge gaps and outstanding questions regarding the PST:

Fieldwork:

- Improved validation techniques for deep-slab PST results, possibly through investigations and testing at deep-slab avalanches;
- Further testing and analysis of the slope-angle dependence of test results in light of Gauthier's (2007) and McClung's (2008, 2009) conflicting data;
- Investigation of the effect of saw-thickness on test results, particularly in relation to weak layer thickness;

- Further testing in storm-snow weak layers, and in PWLs beneath shallow, soft slabs to improve the understanding of test limitations and to determine when the PST is more or less effective than other tests in thin slabs;
- Further use of the PST to test fracture mechanical theories of slab release (e.g. McClung, 2009a; Heierli et al., 2008b). This could include testing of *wider* columns (> 30 cm) using a cord instead of a saw, particularly to address the effect of plane-stress conditions in the PST compared to plane-strain conditions on an avalanche slope, and possibly resolving some false-stable results;
- Side-by-side testing of standard PSTs and the alternative method proposed by McClung (2009) in which the back (up-slope) end is not isolated from the surrounding snowpack, especially at validation sites;
- Regional validation by elevation and/or aspect, strengthened by further field-testing at regionally representative validation sites.

Analysis:

- Further investigation of slab layering and hardness, and weak layer characteristics affecting test results, particularly in the case of false-stable predictions and the differences between SF and ARR results;
- Further analysis of regional avalanche activity and associated regional weather events, potentially to validate test results during times of low or no regional avalanching;
- Multivariate modelling of slope-scale stability, combining the PST with other variables such as profile properties, slope angle, etc., perhaps similar to the ‘Threshold Sum’ method developed by Schweizer and Jamieson (2007) to identify important factors for assessing slope stability;

- Analysis of temporal variability in PST results on individual PWLs at different spatial scales may reveal whether test results on neighbouring days show less variability later in the buried life of the layer than early on.

REFERENCES CITED

Adams, E., and Brown, R. (1983). Metamorphism of dry snow as a result of temperature gradient and vapour density differences. *Annals of Glaciology,* 4, pp. 3-9.

Armstong, R. (1985). *Metamorphism in a subfreezing, seasonal snow cover: the role of thermal and water vapour pressure conditions.* PhD Thesis. University of Colorado, Boulder. 175 pp.

Avalanche.org. (2009). *Avalanche Accidents Database.* http://avalanche.org/accidents.php (last accessed 24 August, 2009)

Bader, H., and Salm, B. (1990). On the mechanics of snow slab release. *Cold Regions Science and Technology*, 17, pp. 287-300.

Bakermans, L. (2006). *Near-Surface Snow Temperature Changes Over Terrain*. MSc Thesis. Dept. of Civil Engineering, University of Calgary, Alberta. 178 pp.

Bažant, Z., Zi, G., and McClung, D. (2003). Size effect law and fracture mechanics of the triggering of dry snow slab avalanches. *Journal of Geophysical Research*, 108(B2), pp. 2119-2129. doi:10.1029/2002JB001884.

Birkeland, K., and Chabot, D. (2006). Minimizing "false-stable" stability test results: Why digging more snowpits is a good idea. *Proceedings of the 2006 International Snow Science Workshop,* Telluride, Colorado, pp. 498-504.

Birkeland, K., and Johnson, R. (1999). The stuffblock snow stability test: comparability with the rutschblock, usefulness in different snow climates, and repeatability between observers. *Cold Regions Science and Technology*, 30, pp. 115-123.

Birkeland, K., and Simenhois, R. (2008). The Extended Column Test: test effectiveness, spatial variability, and comparison with the Propagation Saw Test. *Proceedings of the 2008 International Snow Science Workshop,* Whistler, BC, pp. 401-407

Birkeland, K., Johnson, R., and Herzberg, D. (1996). *The Stuffblock snow stability test.* U.S. Forest Service Missoula Technology and Development Center publication 9623-2836-MTDC.

Birkeland, K., Johnson, R., and Schmidt, D. (1998). Near-surface faceted crystals formed by diurnal recrystalization: a case study of weak layer formation in the mountain snowpack and it's contribution to snow avalanches. *Arctic and Alpine Research,* 30(2), pp. 200-204.

Bradley, C., and Bowles, D. (1967). Strength-load ratio: an index of deep slab avalanche conditions. *Proceedings of the 1966 International Conference on Low Temperature Science,* Sapporo, Japan, pp. 1243-1253.

Breyfogle, S. (1986). *Growth Characteristics of Hoarfrost with Respect to Avalanche Occurence.* Technical Report. Washington State Department of Transportation, Olympia, WA. 40 pp.

Brown, C., and Jamieson, B. (2006). Evolving shear strength, stability, and snowpack properties in storm snow. *Proceedings of the 2006 International Snow Science Workshop,* Telluride, Colorado, pp. 15-21.

Brown, R., Lang, T., St. Lawrence, W., Bradley, C. (1973). A failure criterion for snow. *Journal of Geophysical Research*, 78, pp. 4950-4958.

Bucher, E. (1948). *Contribution to the Theoretical Foundations of Avalanche Defense Construction.* Snow, Ice and Permafrost Research Establishment, Translation 18, 99 pp.

Campbell, C. (2004). *Spatial Variability of Slab Stability and Fracture Properties in Avalanche Start Zones*. MSc Thesis. Dept. of Civil Engineering, University of Calgary, Alberta. 248 pp.

Campbell, C., and Jamieson, B. (2007). Spatial variability of slab stability and fracture characteristics within avalanche start zones. *Cold Regions Science and Technology,* 47, pp. 134-147.

Campbell, C., Bakermans, L., Jamieson, B., and Stethem, C. (2008). Avalanche threats and mitigation measures in Canada. *Proceedings of the 2008 International Snow Science Workshop*, Whistler, BC, pp. 836-844.

Canadian Avalanche Association. (2007). *Observation Guidelines and Recording Standards for Weather, Snowpack and Avalanches.* Revelstoke, BC. Canadian Avalanche Association.

Chabot, D., Kahrl, M., Birkeland, K., Anker, C. (2004). SNOWPILOT: A "new school" tool for collecting, graphing, and databasing snowpit and avalanche occurrence data with a PDA.

Proceedings of the 2004 International Snow Science Workshop, Jackson Hole, Wyoming, pp. 476

Colbeck, S. (1991). The layered characteristics of snow covers. *Reviews of Geophysics,* 29(1), pp. 88-96.

Colbeck, S., and Jamieson, B. (2001). The formation of faceted layers above crusts. *Cold Regions Science and Technology,* 33, pp. 247-252.

Colbeck, S., Akitaya, E., Armstrong, R., Gubler, H., Lafeuille, J., Lied, K., McClung, D., and Morris, E. (1990). *International Classification for Seasonal Snow on the Ground.* International Commission for Snow and Ice (IAHS), World Data Center for Glaciology, University of Colorado. Boulder.

Doswell, C., and Flueck, J. (1989). Forecasting and verifying in a field research project: DOPLIGHT ’87. *Weather and Forecasting*, 4(2), pp. 97-109.

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

Exner, T. (in preparation). *Field Studies of Snowpack Stress and Deformation due to Surface Loads and Temperature Effects.* PhD Thesis. University of Calgary, Alberta.

Fierz, C. (1998). Field observations and weak layer modelling of weak-layer evolution. *Annals of Glaciology,* 26, pp. 7-13.

Föhn, P. (1987). The Rutschblock as a practical tool for slope stability evaluation. *Avalanche Formation, Movements and Effects*. IAHS publication no. 162, Davos, Switzerland, pp. 195-211.

Gauthier, D. (2007). *A Practical Field Test for Fracture Propogation and Arrest in Weak Snowpack Layers in Relation to Slab Avalanche Release.* PhD Thesis. Dept. of Civil Engineering, University of Calgary, Alberta. 302 pp.

Gauthier, D., and Jamieson, B. (2006a). Towards a field test for fracture propagation propensity in weak snowpack layers. *Journal of Glaciology,* 52, pp. 164-168.

Gauthier, D., and Jamieson, B. (2006b). Understanding the propagation of fractures and failures leading to dangerous and destructive snow avalanches: recent developments. *Proceedings of the 2006 Annual Conference of the Canadian Society for Civil Engineering, First Specialty Conference on Disaster Mitigation,* Calgary, Alberta, pp. 010-1–010-9.

Gauthier, D., and Jamieson, B. (2007a). Evaluation of a prototype field test for fracture and failure propagation propensity in weak snowpack layers. *Cold Regions Science and Technology,* 51, pp. 87-97.

Gauthier, D., and Jamieson, B. (2007b). The Propagation Saw Test (PST). *avalanche.ca - The Journal of Canada's Avalanche Community,* 83, pp. 60-63.

Gauthier, D., and Jamieson, B. (2008a). Fracture propagation propensity in relation to snow slab avalanche release: Validating the Propagation Saw Test. *Geophysical Research Letters,* 35, L13501, doi:10.1029/GL034245.

Gauthier, D., and Jamieson, B. (2008b). Predictions of the Propagation Saw Test: comparisons with other instability tests at skier tested slopes. *Proceedings of the 2008 International Snow Science Workshop,* Whistler, BC, pp. 408-414.

Gauthier, D., Ross, C., and Jamieson, B. (2008). Validation of the Propagation Saw Test near whumpfs and avalanches. *Proceedings of the 2008 International Snow Science Workshop,* Whistler, BC, pp. 16-21.

Gubler, H., and Bader, H. (1989). A model of initial failure in slab-avalanche release. *Annals of Glaciology*, 13, pp. 90-95.

Habermann, M., Schweizer, J., and Jamieson, B. (2008). Influence of snowpack layering on human-triggered snow slab avalanche release. *Cold Regions Science and Technology,* 54, pp. 176-182.

Hägeli, P., and McClung, D. (2003). Avalanche characteristics of a transititional snow climate – Columbia Mountains, British Columbia, Canada. *Cold Regions Science and Technology,* 37, pp. 255-276.

Hägeli, P., and McClung D. (2004). Heirarchy theory as a conceptual framework for scale issues in avalanche forecast modeling. *Annals of Glaciology*, 38, pp. 209-214.

Hägeli, P., and McClung, D. (2007). Expanding the snow-climate classification with avalanche-relevent information: initial description of avalanche winter regimes for southwestern Canada. *Journal of Glaciology*, 53, pp. 266-276.

Hägeli, P., McCammon, I., Jamieson, B., Israelson, C., and Statham, G. (2006). The Avaluator – A Canadian rule-based avalanche decision support tool for amateur receationists. *Proceedings of the 2006 International Snow Science Workshop,* Telluride, Colorado, pp. 254-263.

Heierli, J. (2005). Solitary fracture waves in metastable snow stratifications. *Journal of Geophysical Research*, 110. F02008, doi:10.1029/2004JF000178.

Heierli, J. (2009). *Anticrack Model for Slab Avalanche Release.* PhD Thesis. Universität Karlsruhe (TH), Germany. 102 pp.

Heierli, J., and Zaiser, M. (2006). An analytical model for fracture nucleation in collapsible stratifications. *Geophysical Research Letters,* 33. L06501, doi:10.1029/2005GL025311

Heierli, J., and Zaiser, M. (2007). Failure initiation in snow stratifications containing weak layers: Nucleation of whumpfs and slab avalanches. *Cold Regions Science and Technology,* 52, pp. 385-400.

Heierli, J., Gumbsch, P., and Zaiser, M. (2008a). Anticrack nucleation as triggering mechanism for snow slab avalanches. *Science Magazine*, 321, pp. 240-243.

Heierli, J., van Herwijnen, A., Gumbsch, P., and Zaiser, M. (2008b). Anticracks: A new theory of fracture initiation and fracture propagation in snow. *Proceedings of the 2008 International Snow Science Workshop,* Whistler, BC, pp. 9-15.

Hendrikx, J., Birkeland, K. (2008). Slope scale spatial variability across time and space: comparison of results from two different snow climates. *Proceedings of the 2008 International Snow Science Workshop,* Whistler, BC, pp. 155-162.

Hendrikx, J., Birkeland, K., Clark, M. (2009). Assessing changes in the spatial variability of the snowpack fracture propagation propensity over time. *Cold Regions Science and Technology*, 56, pp. 152-160.

Jaccard, C. (1966). Stabilité des plaques de neige. *Proceedings of the International Symposium on scientific aspects of snow and ice avalanches,* Davos, Switzerland. IAHS publication no. 69, pp. 170-181.

Jamieson, B. (1995). *Avalanche Prediction for Persistent Snow Slabs.* PhD Thesis. Dept. of Civil Engineering, University of Calgary, Alberta. 255 pp.

Jamieson, B. (1999). The Compression Test - after 25 years. *The Avalanche Review,* 18(1), pp. 10-12.

Jamieson, B. (2003). Fracture propagation and resistance in weak snowpack layers. *Avalanche News,* 67.

Jamieson, B. (2006). Formation of refrozen snowpack layers and their role in slab avalanche release. *Reviews of Geophysics,* 44, RG2001, doi:10.1029/2005RG000176.

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

Jamieson, B., and Johnston, C. (1993a). Experience with Rutshcblocks. *Proceedings of the 1992 International Snow Science Workshop*, Brekenridge, Colorado, pp. 150-159.

Jamieson, B., and Johnston, C. (1993b). Shear frame stability parameters for large-scale avalanche forecasting. *Annals of Glaciology,* 18, pp. 268-273.

Jamieson, B., and McClung, D. (1996). Terminology: whumpfs and propagating shear fractures. *Avalanche News,* 49, pp. 19

Jamieson, B., and Schweizer, J. (2000). Texture and strength changes of buried surface hoar layers with implications for dry snow-slab avalanche release. *Journal of Glaciology,* 46, pp. 151-160.

Jamieson, B., and Schweizer, J. (2005). Using a checklist to assess manual snow profiles. *Avalanche News,* 72, pp. 56-59.

Jamieson, B., Geldsetzer, T., and Stethem, C. (1998). Forecasting for deep slab avalanches. *Cold Regions Science and Technology,* 33, pp. 275-290.

Jamieson, B., Zeidler, A., and Brown, C. (2007). Explanation and limitations of study plot stability indices for forecasting dry snow slab avalanches in surrounding terrain. *Cold Regions Science and Technology,* 50, pp. 23-34.

Jamieson, B., Campbell, C., and Jones, A. (2008). Verification of Canadian avalanche bulletins including spatial and temporal scale effects. *Cold Regions Science and Technology,* 51, pp. 204-213.

Johnson, B. C. (2000). *Remotely Triggered Slab Avalanches.* MSc Thesis. Dept. of Civil Engineering, University of Calgary, Alberta. 98 pp.

Johnson, B. C., Jamieson, B., Stewart, R. (2004). Seismic measurements of fracture speed in a weak snowpack layer. *Cold Regions Science and Technology*, 40, pp. 21-45.

Johnson, G. T. (2000). *Observations of Faceted Crystals in Alpine Snowpacks*. MSc Thesis. Dept. of Civil Engineering, University of Calgary, Alberta. 100 pp.

Johnson, R., and Birkeland, K. (1998). Effectively using and interpreting stability tests. *Proceedings of the 1998 International Snow Science Workshop,* Sunriver, Oregon, pp. 562-565.

Johnson, R., and Birkeland, K. (2002). Integrating shear quality into stability test results. *Proceedings of the 2002 International Snow Science Workshop,* Penticton, BC, pp. 508-513.

Kronholm, K. (2004). *Spatial Variability of Snow Mechanical Properties with regard to Avalanche Formation*. PhD Thesis. University of Zurich, Switzerland. 187 pp.

LaChapelle, E., and Armstrong, R. (1977). Temperature patterns in an alpine snow cover and their influence on snow metamorphism. U.S. Army Research Office technical report, Institute of Arctic and Alpine Research, University of Colorado, Boulder. (not seen).

Lackinger, B. (1989). Supporting forces and stability of snowslab avalanches: a parameter study. *Annals of Glaciology*, 13, pp. 140-145.

Landry, C. (2002). *Spatial Variations in Snow Stability on Uniform Slopes: Implications for Extrapolation to Surrounding Terrain.* MSc Thesis. Montana State University, Bozeman, Montana. 194 pp.

Lang, R., Leo, B., and Brown, R. (1984). Observations on the growth process and strength characteristics for surface hoar. *Proceedings of the 1984 International Snow Science Workshop,* Aspen, Colorado, pp. 188-195.

Lang, T., and Brown, R. (1975). Stress concentrations in a sloping snowpack from geometric imperfections. IAHS publication no. 114, pp. 311-320.

Louchet, F. (2001a). Creep instability of the weak layer and natural slab avalanche triggerings. *Cold Regions Science and Technology*, 33, pp. 141-146.

Louchet, F. (2001b). A transition in dry-snow slab avalanche triggering modes. *Annals of Glaciology*, 32, pp. 285-289.

Mann, H., and Whitney, D. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Annals of Mathematical Statistics,* 18(1), pp. 50-60.

McCammon, I., and Schweizer, J. (2002). A field method for identifying structural weakness in the snowpack. *Proceedings of the 2002 International Snow Science Workshop*, Penticton, BC, pp 477-481.

McClung, D., (1977). Direct simple shear tests on snow and their relation to slab avalanche formation. *Journal of Glaciology*, 19(81), pp. 101-109.

McClung, D. (1979). Shear fracture precipitated by strain softening as a mechanism of dry slab avalanche release. *Journal of Geophysical Research,* 84(B7) , pp. 3519-3526.

McClung, D. (1981). Fracture mechanical models of dry slab avalanche release. *Journal of Geophysical Research*, 86(B11) pp. 10783-10790.

McClung, D. (1987). Mechanics of snow slab failure from a geotechnical perspective. *Avalanche Formation, Movements and Effects*. IAHS publication no. 162, Davos, Switzerland, pp. 475-508.

McClung, D., (2000). Predictions in avalanche forecasting. *Annals of Glaciology,* 31, pp. 377-381.

McClung, D. (2002). Fracture toughness for dry slab avalanches. *Proceedings of the 2002 International Snow Science Workshop*, Penticton, BC.

McClung, D. (2005a). Dry slab avalanche shear fracture properties from field measurements. *Journal of Geophysical Research*, 110, F04005, doi:10.1029/2005JF000291

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

McClung, D. (2007a). Fracture energy applicable to dry snow slab avalanche release. *Geophysical Research Letters*, 34, L02503, doi:10.1029/2006GL028238.

McClung, D. (2007b). Dry snow slab shear fracture speeds. *Geophysical Research Letters*, 34, L10502, doi:10.1029/2007GL029261.

McClung, D. (2008). Avalanche fracture mechanics results for practitioners (abstract). *Proceedings of the 2008 International Snow Science Workshop*, Whistler, BC, pp. 25.

McClung, D. (2009a). Dry snow slab quasi-brittle fracture initiation and verification from field tests. *Journal of Geophysical Research*, 114, F01022, doi:10.1029/2007JF000913.

McClung, D. (2009b). Dimensions of dry snow slab avalanches from field measurements. *Journal of Geophysical Research*, 114, F01006, doi:10.1029/2007JF000941.

McClung, D., and Schaerer, P. (2006). *The Avalanche Handbook.* Seattle, WA: The Mountaineers Books.

McClung, D, and Schweizer, J. (2006). Fracture toughness of dry snow slab avalanches from field measurements. *Journal of Geophysical Research*, 111, F04008, doi:10.1029/2005JF000403.

Miller, D., Adams, E., and Brown, R. (2003). A microstructural approach to predict dry snow metamorphism in generalized thermal conditions. *Cold Regions Science and Technology,* 37(3), pp. 213-226.

Moner, I., Gavaldà, J., Bacardit, M., Garcia, C., and Martī, G. (2008). Application of field stability evaluation methods to the snow conditions of the Eastern Pyrenees. *Proceedings of the 2008 International Snow Science Workshop*, Whistler, BC, pp. 386-392.

Moore, M. (1982). Temperature gradient weakening near crusts in a maritime snowpack. *Proceedings of the 1982 International Snow Science Workshop.* Bozeman, Montana.

Nye, J. (1975). General discussion, *Proceedings of the Symposium at Grindelwald 1974 – Snow Mechanics,* IAHS publication no. 114, pp. 442. (not seen)

Perla, R. (1969). Strength test on newly fallen snow. USDA Forest Service Res Note RM-150, 12 pp.

Perla, R. (1980). Avalanche release, motion and impact. In S. Colbeck (Ed.), *Dynamics of Snow and Ice Masses.* New York: Academic Press. pp. 397-462.

Perla, R., and LaChapelle, E. (1970). A theory of snow slab failure. *Journal of Glaciology*, 75, pp. 7619-7627.

Ross, C., and Jamieson, B. (2008). Comparing fracture propagation tests and relating test results to snowpack characteristics. *Proceedings of the 2008 International Snow Science Workshop,* Whistler, BC, pp. 376-385.

Ross, S., and Jamieson, B. (2009). Validating the Propagation Saw Test on the slope scale. *Proceedings of the 2009 International Snow Science Workshop*, Davos, Switzerland.

Schweizer, J. (1999). Review of dry snow slab avalanche release. *Cold Regions Science and Technology,* 30, pp. 43-57.

Schweizer, J. (2002). The Rutschblock test – Procedure and application in Switzerland. *The Avalanche Review,* 20, pp. 14-15.

Schweizer, J. and Camponovo, C. (2001). The skier's zone of influence in triggering slab avalanches. *Annals of Glaciology,* 32, pp. 314-320.

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

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

Schweizer, J., and Jamieson, B. (2007). A threshold sum approach to stability evaluation of manual snow profiles. *Cold Regions Science and Technology,* 47, pp. 50-59

Schweizer, J., and Jamieson, B. (2009). Snowpack tests for assessing snow-slope instability. *Annals of Glaciology*, 54 (in press).

Schweizer, J., and Wiesinger, T. (2001). Snow profile interpretation for stability evaluation. *Cold Regions Science and Technology*, 33, pp. 179-188.

Schweizer, J., Camponovo, C., Fierz, C., and Föhn, P. (1995a). Skier triggered slab avalanche release – some practical implications. *Proceedings of the International Symposium: Science and Mountain – the contribution of scientific research to snow, ice and avalanche safety,* Chamonix, France, pp. 309-315.

Schweizer, J., Schneebeli, M., Fierz, C., and Föhn, P. (1995b). Snow mechanics and avalanche formation: Field experiments on the dynamic response of the snow cover. *Surveys in Geophysics,* 16, pp. 621-633.

Schweizer, J., Jamieson, B., and Schneebeli, M. (2003a). Snow Avalanche Formation. *Reviews of Geophysics,* 41(4), pp. 1016.

Schweizer, J., Kronholm, K., and Wiesinger, T. (2003b). Verification of regional snowpack stability and avalanche danger. *Cold Regions Science and Technology,* 37, pp. 277-288.

Schweizer, J., Fierz, C., and Jamieson, B. (2004). Assessing the probability of skier triggering from snow layer properties. *Proceedings of the 2004 International Snow Science Workshop,* Jackson Hole, Wyoming. pp. 192-198

Schweizer, J., McCammon, I., and Jamieson, B. (2006). Snow slope stability evaluation using concepts of fracture mechanics. *Proceedings of the 2006 International Snow Science Workshop*, Telluried, Colorado, pp. 211-218.

Schweizer, J., Kronholm, K., Jamieson, B., and Birkeland, K. (2008). Review of spatial variability of snowpack properties and its importance for avalanche formation. *Cold Regions Science and Technology,* 51, pp. 253-272.

Schweizer, J., Mitterer, C., and Stoffel, L. (2009). On forecasting large and infrequent snow avalanches. *Cold Regions Science and Technology* (in press).

Sigrist, C. (2006). *Measurement of Fracture Mechanical Properties of Snow and Application to Dry Snow Slab Avalanche Release.* PhD Thesis, Swiss Federal Institute of Technology, Zurich. 139pp.

Sigrist, C., and Schweizer, J. (2007). Critical energy release rates of weak snowpack layers determined in field experiments. *Geophysical Research Letters*, 34. L03502, doi:10.1029/2006GL028576.

Simenhois, R., and Birkeland, K. (2006). The Extended Column Test: a field test for fracture initiation and propagation. *Proceedings of the 2006 International Snow Science Workshop,* Telluride, Colorado, pp. 79-85.

Simenhois, R., and Birkeland, K. (2007). An update on the Extended Column Test: new recording standards and additional data analyses. *The Avalanche Review,* 26(2), pp. 17-18.

Simenhois, R., and Birkeland, K. (2008a). The effect of changing slab thickness on fracture propagation. *Proceedings of the 2008 International Snow Science Workshop*, Whistler, BC, pp. 755-760.

Simenhois, R., and Birkeland, K. (2008b). Extended Column Test results during surface warming: Does surface warming affect fracture propagation potential? *Proceedings of the 2008 International Snow Science Workshop*, Whistler, BC, pp. 36-35.

Simenhois, R., and Birkeland, K. (2009). The Extended Column Test: Test effectiveness, spatial variability, and comparison with the Propagation Saw Test. *Cold Regions Science and Technology*, 59, pp. 210-216.

Smith, F., and Curtis, J. (1975). Stress analysis and failure prediction in avalanche snowpacks. IAHS Publication no. 114, pp. 332-340.

Statham, G., McMahon, B., and Tomm, I. (2006). The Avalanche Terrain Exposure Scale. *Proceedings of the 2006 International Snow Science Workshop*, Telluride, Colorado, pp. 491-497.

Stewart, W. (2002). *Spatial Variability of Slab Stability within Avalanche Starting Zones.* MSc Thesis. University of Calgary, Alberta. 100 pp.

van Herwijnen, A. (2005). *Fractures in Weak Snowpack Layers in Relation to Slab Avalanche Release*. PhD Thesis. Dept. of Civil Engineering, University of Calgary, Alberta. 269 pp.

van Herwijnen, A., and Jamieson, B. (2002). Interpreting fracture character in stability tests. *Proceedings of the 2002 International Snow Science Workshop,* Penticton, BC, pp. 514-520.

van Herwijnen, A., and Jamieson, B. (2003). An update on fracture character in stability tests. *Avalanche News*, 66, pp. 26-28.

van Herwijnen, A., and Jamieson, B. (2004a). Fracture character in compression tests. *Proceedings of the 2004 International Snow Science Workshop,* Jackson Hole, Wyoming, pp. 182-191.

van Herwijnen, A., and Jamieson, B. (2004b). More results on fracture characterization in compression tests. *Avalanche News*, 68, pp. 38-41.

van Herwijnen, A., and Jamieson, B. (2005). High-speed photography of fractures in weak snowpack layers. *Cold Regions Science and Technology,* 43, pp. 71-82.

van Herwijnen, A., and Jamieson, B. (2007a). Fracture character in compression tests. *Cold Regions Science and Technology,* 47, pp. 60-68

van Herwijnen, A., and Jamieson, B. (2007b). Snowpack properties associated with fracture initiation and propagation resulting in skier-triggered dry snow avalanches. *Cold Regions Science and Technology,* 50, pp. 13-22.

van Herwijnen, A., Heierli, J., and Schweizer, J. (2008). Field study on fracture propagation in weak snowpack layers. *Proceedings of the 2008 International Snow Science Workshop*, Whistler, BC, pp. 1-8.

van Herwijnen, A., Bellaire, S., and Schweizer, J. (2009). Comparison of micro-structural snowpack parameters derived from penetration resistane measurements with fracture character observations from compression tests. *Cold Regions Science and Technology*, 59, pp. 193-201.

Winkler, K., and Schweizer, J. (2008). Comparison of different snow stability tests including the Extended Column Test. *Proceedings of the 2008 International Snow Science Workshop,* Whistler, BC, pp. 393-400.

Winkler, K., and Schweizer, J. (2009). Comparison of snow stability tests: Extended column test, rutschblock test and compression test. *Cold Regions Science and Technology*, 59, pp. 217-226.

APPENDIX A – Slope-scale Validation Data

The slope-scale validation dataset (2007-2009) analysed in Section 4.2 is presented in Appendix A, showing all the site-layers of confirmed propagation or confirmed initiation without propagation, and all the associated PSTs. ECT results at validation sites in 2008 are also included. Day-old avalanche sites are identified in italics.

NOTATION:

HST	storm snow
Hr	helicopter-remote (trigger)
L	likely
Na	natural avalanche (trigger)
NR	No Result
Sa	skier-accidental (trigger)
Sc	ski cut
Sr	skier-remote (trigger)
Sz	avalanche size class (1-5)
UL	unlikely
Xe/Xr	explosive/explosive-remote (trigger)

A.1 2007 slope-scale validation sites and PST results

Test Date	PWL	Layer Depth	Slope	Standard PST		PST % Cut to END 1	2	3	4	Sc Result
06-Jan-07	unk	6-12	38	1	UL	93%				Sc NR
19-Jan-07	Jan 18		34	1	UL (SF)	SF				Sr
02-Feb-07	windslab		38	3	2L/1UL	40%	ARR	20%		Sa/Sc
06-Feb-07	Feb 4 SH		33	2	2UL	ARR	88%			Sa/Sc
06-Feb-07	Feb 4 SH	~ 13	35	2	UL	SF	SF			Sc NR
08-Feb-07	Feb 4 SH		32	1	UL	SF				Sc NR
09-Feb-07	Feb 4 SH		37	3	UL	SF	SF	SF		Sc NR
10-Feb-07	Feb 4 SH		33	2	UL	SF	ARR			Sc NR
12-Feb-07	Feb 4 SH		27	3	3L	26%	22%	25%		Sa/Sc
12-Feb-07	unk			1	L	25%				whumpf
13-Feb-07	windslab	10-25	44	2	1UL/1L	61%	47%			Sa/Sc
14-Feb-07	Feb 4 SH		29	1	L	30%				Sa/Sc
14-Feb-07	Feb 4 SH		24	3	3L	27%	36%	30%		Sa/Sc
16-Feb-07	Feb 4 SH		45	2	1UL/1L	50%	48%			whumpf
17-Feb-07	Feb 4 SH		21	1	L	39%				Sr
19-Feb-07	Feb 4 SH		35	1	L	43%				whumpf
21-Feb-07	Feb 4 SH		26	2	L	26%	33%			whumpf
23-Feb-07	Feb 4 SH		24	1	L	29%				whumpf
24-Feb-07	Feb 4 SH		30-37	1	0.5	51%				Sa/Sc
25-Feb-07	Feb 4 SH		19	1	L	39%				whumpf
28-Feb-07	Feb 4 SH		27	2	L	29%	39%			Sr
01-Mar-07	Feb 4 SH		29	2	L	28%	29%			whumpf
				38	22 sites	and	22	validation site-layers		

A.2 2008 slope-scale validation sites and PST and ECT results

Test Date	PWL	Layer Depth	Slope	Standard PST		PST % Cut if to END Test Number: 1	2	3	4	Sc Result	ECT 1	ECT 2
2-Feb-08	Jan 26 SH	45-50	25	2	UL/UL	ARR	ARR			Sc 0.5	ECTN 12	ECTP 14
4-Feb-08	*Jan 26 SH*	*39*	*38*	*2*	*UL/UL*	*77%*	*75%*			*day-old Sa*	*ECTP 15*	*ECTP 15*
22-Feb-08	*Jan 26 SH*	*57*	*37*	*2*	*L/L*	*46%*	*43%*			*day-old Sr*	*ECTP 29*	*ECTNR*
2-Mar-08	Feb 23 SH	30-35	48	2	L/L	23%	23%			Sr	ECTP 13	ECTP 13
3-Mar-08	Feb 23 SH	40	32	4	UL/L/L/UL	55%	35%	21%	SF	near Xe/Xr	ECTP 12	ECTP 12
6-Mar-08	*Feb 23 SH*	*60-64*	*34*	*3*	*L/L/L*	*43%*	*37%*	*41%*		*day-old Sr*	*ECTP 22*	*ECTP 22*
6-Mar-08	Feb 23 SH	30	35	2	L/L	29%	33%			wumph	ECTP 12	ECTP 12
7-Mar-08	Feb 23 SH	35-40	21-29	2	L/L	34%	30%			Sc	ECTP 13	ECTP 14
10-Mar-08	Feb 23 SH	43	20	2	L/L	27%	26%			wumph	ECTP 14	ECTP 12
12-Mar-08	Feb 26 SH	38-40	28	2	L/L	32%	35%			Na's	ECTP 20	ECTP 21
19-Mar-08	*Feb 23 SH*	*90*	*20*	*2*	*.5/.5*	*50%*	*53%*			*day-old Xe/Sr*	*ECTNR*	*ECTNR*
19-Mar-08	*Mar 9 SH*	*53*	*20*	*2*	*UL/UL*	*94%*	*100%*			*day-old Xe/Sr*	*ECTN 26*	*ECTNR*
21-Mar-08	Feb 23 SH	80	16	1	L	47%				wumph	ECTNR	~
21-Mar-08	Mar 9 SH	48	16	1	L	38%	47%*			wumph	ECTP 18	~
4-Feb-08	Jan 26 SH	65	29	2	L/UL	29%	63%			Sc	ECTP 27	ECTP 25
07-Feb-08	Jan 26 SH	65	38	2	L/L	29%	34%			whumpf	ECTN 30	ECTP 23
08-Feb-08	*Jan 26 SH*	*89*	*40*	*2*	*UL/UL*	*83%*	*73%*			*day-old Sa*	*ECTNR*	*ECTP 21*
28-Feb-08	*??*	*38*	*36*	*2*	*UL/UL*	*ARR*	*SF*			*day-old Na*	*ECTN 18*	*ECTN 20*
03-Mar-08	Feb 26 SH	52	43	2	.5/.5	47%	50%			Sc	ECTN 27	ECTP 23
04-Mar-08	Feb 26 SH	79	37	3	L/L/L	34%	30%	31%		Sc	ECTNR	ECTN 21
05-Mar-08	Feb 26 SH	37	37	2	L/L	32%	30%			Sc	ECTN 18	ECTP 12
10-Mar-08	Feb 26 SH	49	38	2	L/L	39%	40%			Sc	ECTP 11	ECTP 13
13-Mar-08	*suncrust*	*40*	*44*	*2*	*L/L*	*45%*	*45%*			*day-old Hr*	~	~
13-Mar-08	*Feb 26 SH*	*34*	*37*	*2*	*SF/SF*	*SF*	*SF*			*day-old Hr*	~	~

Continued...

A.2 Continued...

Test Date	PWL	Layer Depth	Slope	Standard PST		Test Number: 1	2	3	4	Sc Result	ECT 1	ECT 2
18-Mar-08	Mar 8	55	29	2	UL/UL	90%	81%			Sc	ECTNR	ECTNR
19-Mar-08	Mar 8	34	38	1	UL	ARR				Sc	ECTNR	ECTNR
20-Mar-08	*Feb 26 SH*	*71*	*42*	*2*	*UL/UL*	*85%*	*ARR*			*day-old Sa*	*ECTNR*	*ECTNR*
21-Mar-08	Mar 16	52	38	2	UL/UL	98%	95%			Sc	ECTNR	ECTNR
				57	25 sites	and	28	validation site-layers				

A.3 2009 slope-scale validation sites and PST results

Test Date	PWL	Layer Depth	Slope	Standard PST		PST % Cut if to END Test Number: 1	2	3	4	Sc Result
03/01/2009	25/12/2008	55	32	2	1UL/1L	57%	38%			whumpf
11/01/2009	HST	30	25	2	UL (ARR)	ARR	ARR			Sc in HST
02/02/2009	HST	12	22	2	UL (SF)	SF	SF			Sr/whumpfs
04/02/2009	27/01/2009	38	28	2	L	34%	42%			whumpf
05/02/2009	27/01/2009	53	28	2	L	30%	26%			whumpf
07/02/2009	27/01/2009	52	5	3	L	43%	33%	38%		whumpf
09/02/2009	27/01/2009	43	30	2	L	42%	43%			whumpf
11/02/2009	27/01/2009	38	39	2	L	36%	38%			Sc/Sa
11/02/2009	unk	25	39	2	1SF/1ARR	SF	ARR			Sc
13/02/2009	unk	20	30	2	1UL/1SF	79%	SF			ScNR
23/02/2009	27/01/2009	31	24	3	2L/1UL (SF)	42%	37%	SF		ScNR
25/02/2009	22/02/2009	47	36	2	UL (ARR)	ARR	ARR			ScNR
27/02/2009	27/01/2009	77	42	2	L	29%	39%			Sc
27/02/2009	22/02/2009	41	42	2	UL (ARR)	ARR	ARR			Sc

Continued...

A.3 Continued...

Test Date	PWL	Layer Depth	Slope	Standard PST		Test Number: 1	2	3	4	Sc Result
27/02/2009	27/01/2009	60	30	3	L	39%	39%	39%		Sc
27/02/2009	22/02/2009	34	30	2	UL (ARR)	ARR	ARR			Sc
01/03/2009	22/02/2009	46	28	2	UL(ARR)	ARR	ARR			Sc
11/03/2009	01/03/2009	41	31	2	UL (SF)	SF	SF			Sc
15/03/2009	unk	69	32	2	L	37%	39%			whumpf
15/03/2009	01/03/2009	33	32	2	L	43%	22%			whumpf
15/03/2009	unk	18	32	2	UL(SF)	SF	SF			Sc
24/03/2009	01/03/2009	22	32	2	UL	100%	97%			Sc
25/03/2009	HST/crust	15	24	2	UL (SF/ARR)	SF	ARR			Sc
28/01/2009	27/01/2009	23.5	31	1	UL(SF)	SF				Sc NR
28/01/2009	lower crust	33	31	2	UL(SF)	SF	SF			Sc NR
29/01/2009	27/01/2009	35	28	2	UL(SF)	SF	SF			Sc NR
02/02/2009	unk	20	40	1	UL(SF)	SF				Sc
09/02/2009	unk	24	4	2	1L/1UL(ARR)	25%	ARR	0%	0%	whumpf
10/02/2009	unk	51	23	2	L	30%	25%	0%	0%	whumpf/Sr
26/02/2009	22/02/2009	50	10	4	2L/2UL(1ARR)	22%	27%	89%	ARR	Sc
13/03/2009	*unk*	*34*	*37*	*2*	*UL(SF)*	*SF*	*SF*			*day-old Hr 1.5*
17/03/2009	unk	22	31	2	UL	86%	87%	0%	0%	sc NR
18/03/2009	*27/01/2009*	*66*	*37*	*4*	*1L/3UL(ARR)*	*30%*	*ARR*	*ARR*	*ARR*	*day-old Sa 1.5*
20/03/2009	27/01/2009	91	10	3	L	42%	46%	34%		whumpf

28 sites	and	34	validation site-layers

APPENDIX B – Raw Short-Scaled PST Data

			Short-Scaled PST % Cut			
WL depth	Isolated Length	Cut Length	END	ARR	SF	PIVOT
55	57	32	56%			
55	54	28	52%			
62	62	28	45%			
62	63	28	44%			
71	70	33	47%			
73	73	31		42%		
72	76	33		43%		
33	33	7				21%
30	30	8				27%
30	30	9				30%
69	70	26	37%			
68	71	33	46%			
55	55	33	60%			
58	62	31	50%			
47	47	15		32%		
43	43	12		28%		
13	13	7	54%			
13	13	5	38%			
70	72	24	33%			
68	71	25	35%			
46	46	12		26%		
70	66	32	48%			
64	66	30	45%			
46	46	15	33%			
47	46	17	37%			
65	66	35	53%			
46	44.5	17		38%		
67	67	33	49%			
47	48	16				33%
45	46	17	37%			
27	27	10		37%		
45	47	25	53%			
30	32	16		50%		

Continued...

WL Depth	Isolated Length	Cut Length	END	ARR	SF	PIVOT
36	35	14	40%			
35	35	14	40%			
23	23	11	48%			
23	23	12	52%			
53	54	26	48%			
53	53	26	49%			
31	42	22	52%			
34	34	18	53%			
32	32	17	53%			
38	38	20	53%			
39	39	16	41%			
41	41	17	41%			
52	53	26	49%			
51	51	38	75%			
52	52	30	58%			
54	54	54	100%			
51	51	51	100%			
37	38	22	58%			
38	38	22	58%			
42	42	15	36%			
42	44	19	43%			
49	49	34		69%		
38	40	20	50%			
38	38	23	61%			
26	26	18	69%			
28	28	15	54%			
56	57	30	53%			
40	40	29	73%			
57	57	33	58%			
40	42	29	69%			
62	62	30	48%			
20	24	16	67%			
62	61	27	44%			
21	21	13.5	64%			
31	32	17				53%
31	31	14				45%
31	31	14				45%
31	31	15				48%

Continued...

WL Depth	Isolated Length	Cut Length	END	ARR	SF	PIVOT
69	70	37		53%		
28	28	13	46%			
28	27	13	48%			
67	67	43	64%			
29	29	20	69%			
29	29	13	45%			
74	74	28	38%			
68	68	32	47%			
47	47	23	49%			
77	99	38	38%			
75	79	36	46%			
43	43	24	56%			
73	73	35	48%			
43	43	20	47%			
77	77	27	35%			
41	41	28	68%			
76	76	29	38%			
41	42	32	76%			
42	42	23	55%			
58	58	24	41%			
33	33	21	64%			
63	64	30	47%			
35	37	24		65%		
62	62	40	65%			
46	44	20	45%			
65	65	39	60%			
46	46	32	70%			
29	32	29	91%			
28	27	17	63%			
60	60	23	38%			
39	39	23	59%			
57	60	27	45%			
39	39	25	64%			
40	40	20.5		51%		
84	82	74	90%			
41	42	25	60%			
42	42	25	60%			
42	42	24		57%		

Continued...

WL Depth	Isolated Length	Cut Length	END	ARR	SF	PIVOT
84	86	40	47%			
80	80	37	46%			
28	30	20				67%
29	31	21				68%
32	32	17	53%			
32	33	18	55%			
34	32	18	56%			
34	34	20	59%			
51	51	28	55%			
77	77	30	39%			
47	47	19	40%			
79	79	34	43%			
32	33	10	30%			
32	32	11	34%			
18	18	7	39%			
18	17	7	41%			
15	15	14.8				99%
15	30	20				67%
63	63	44	70%			
43	44	28	64%			
57	56	41	73%			
42	42	25	60%			
23	23	14				61%
23	24	13				54%
12	15	9				60%
13	15	11.5				77%
71	71	36	51%			
61	61	21	34%			
71	71	31	44%			
70	74	29	39%			
61	61	23	38%			
72	72	36	50%			
61	61	31	51%			
10	10	7	70%			
10	10	7	70%			
10	10	7	70%			
90	90	36		40%		
90	87	44		51%		

Continued...

WL Depth	Isolated Length	Cut Length	END	ARR	SF	PIVOT
37	41	14	34%			
88	88	20	23%			
91	89	45	51%			
88	93	39		42%		
25	25	16				64%
30	30	17	57%			
23	30	18				60%
35	37	15	41%			
34	34	14		41%		
35	35	15	43%			
37	36	15	42%			
37	37	10	27%			
37	37	10	27%			
20	32	5	16%			
20	34	6	18%			
42	42	17	40%			
46	45	23	51%			
30	31	13	42%			
30	30	13	43%			
49	52	26	50%			
49	52	22	42%			
62	64	27				42%
63	65	28				43%
25	25	12	48%			
23	22	11	50%			
20	20	9	45%			
51	52	21	40%			
51	51	22	43%			
52	62	37	60%			
52	52	27	52%			
43	43	21	49%			
44	44	18	41%			
22	22	16	73%			
22	23	16	70%			
20	21	16	76%			
60	66	36	55%			
60	60	33.5	56%			
13	13	6.5	50%			

Continued...

WL Depth	Isolated Length	Cut Length	END	ARR	SF	PIVOT
13	13	6.5	50%			
75	80	35	44%			
51	50	20	40%			
75	77	30	39%			
52	50	20	40%			
54	53	15	28%			
28	53	20	38%			
51	47	20	43%			
52	54	51	94%			
82	90	42	47%			
77	76	34	45%			
69	76	37	49%			
64	62	30	48%			
70	76	35	46%			
69	70	39	56%			
42	50	17	34%			
40	48	21	44%			
63	65	28	43%			
56	61	25	41%			
63	60	26		43%		
66	63	26	41%			
64	65	30	46%			
68	65	23	35%	2%		
65	69	25	36%			
60	56	46			82%	
61	62	21			34%	
75	65	35	54%			
53	55	22		40%		
53	56	20		36%		
72	72	47	65%			
70	66	44	67%			
34	30	14	47%			
90	90	26	29%			
91	91	27	30%			
93	96	41	43%			

APPENDIX C – Expert Rating Comparison Data

The comparisons of expert ratings of propagation propensity and PST results analyzed in Section 4.5 are presented in Appendix C. Individual tests and ratings on the same PWL and same day for all eight significant PWLs in 2008 and 2009 are given once they were buried deeper than 70 cm. The column 'INT' refers to the interpretation of test results and the '% Cut' column refers to the percent cut when a fracture propagated END, otherwise SF or ARR is shown. Comparison in layers shallower than 70 cm are not presented; however all other comparisons can be reproduced from the following table.

STANDARD PSTs DEEPER than 70cm vs. EXPERT RATING (INDIVIDUAL TESTS)								Expert Rating			
Test Date	PWL	Layer Depth	PST	INT	% Cut	Site	Veg. Zone	Date	Prop	Applies to	MATCH
08/01/2009	25/12/2008	137	L	L	31	Fidelity	TL	08/01/2009	L	ALP, TL	TRUE
08/01/2009	25/12/2008	137	L	L	47	Fidelity	TL	08/01/2009	L	ALP, TL	TRUE
11/01/2009	25/12/2008	140	L	L	38	Fidelity	TL	11/01/2009	L	ALP, TL	TRUE
11/01/2009	25/12/2008	140	UL	U	ARR	Fidelity	TL	11/01/2009	L	ALP, TL	FALSE
11/01/2009	25/12/2008	140	UL	U	ARR	Fidelity	TL	11/01/2009	L	ALP, TL	FALSE
12/01/2009	25/12/2008	118	L	L	45	Hermit	TL	12/01/2009	L	ALP, TL	TRUE
12/01/2009	25/12/2008	118	UL	U	72	Hermit	TL	12/01/2009	L	ALP, TL	FALSE
12/01/2009	25/12/2008	118	UL	U	ARR	Hermit	TL	12/01/2009	L	ALP, TL	FALSE
13/01/2009	25/12/2008	130	UL	U	ARR	Cheops	TL	13/01/2009	L	ALP, TL	FALSE
13/01/2009	25/12/2008	130	UL	U	ARR	Cheops	TL	13/01/2009	L	ALP, TL	FALSE
24/01/2009	25/12/2008	170	UL	U	ARR(60)	Bruin's Rdg	TL	24/01/2009	E	ALP, TL	FALSE
24/01/2009	25/12/2008	170	UL	U	95	Bruin's Rdg	TL	24/01/2009	E	ALP, TL	FALSE
23/02/2009	25/12/2008	85	UL	U	ARR	Loop Brook	BTL	23/02/2009	U	ALP, TL	TRUE
26/02/2009	27/01/2009	74	L	L	38	Fidelity	TL	26/02/2009	L	TL, BTL	TRUE
26/02/2009	27/01/2009	74	L	L	41	Fidelity	TL	26/02/2009	L	TL, BTL	TRUE
27/02/2009	27/01/2009	77	L	L	29	Fidelity	BTL	27/02/2009	L	TL, BTL	TRUE
27/02/2009	27/01/2009	77	L	L	39	Fidelity	BTL	27/02/2009	L	TL, BTL	TRUE
11/03/2009	22/02/2009	83	UL	U	56	RoundHill	ALP	11/03/2009	L	ALP, TL, BTL	FALSE
11/03/2009	22/02/2009	83	UL	U	51	RoundHill	ALP	11/03/2009	L	ALP, TL, BTL	FALSE
14/03/2009	27/01/2009	101.5	UL	U	ARR(61)	RoundHill	ALP	14/03/2009	E	TL, BTL	FALSE
14/03/2009	27/01/2009	101.5	UL	U	ARR(50)	RoundHill	ALP	14/03/2009	E	TL, BTL	FALSE
14/03/2009	22/02/2009	77.5	UL	E	ARR(42)	RoundHill	ALP	14/03/2009	L	ALP, TL, BTL	FALSE
14/03/2009	22/02/2009	77.5	UL	E	ARR(29)	RoundHill	ALP	14/03/2009	L	ALP, TL, BTL	FALSE

Continued...

Test Date	PWL	Layer Depth	PST	INT	% Cut	Site	Veg. Zone	Date	Prop	Applies to	MATCH
21/03/2009	27/01/2009	117.5	L	L	46	South Run	TL	21/03/2009	U	TL, BTL	FALSE
21/03/2009	27/01/2009	117.5	L	L	34	South Run	TL	21/03/2009	U	TL, BTL	FALSE
21/03/2009	27/01/2009	113.5	L	L	35	South Run top	TL	21/03/2009	U	TL, BTL	FALSE
21/03/2009	27/01/2009	113.5	L	L	45	South Run top	TL	21/03/2009	U	TL, BTL	FALSE
21/03/2009	22/02/2009	90	L	L	42	South Run	TL	21/03/2009	L	ALP, TL, BTL	TRUE
21/03/2009	22/02/2009	90	L	L	50	South Run	TL	21/03/2009	L	ALP, TL, BTL	TRUE
25/03/2009	22/02/2009	108	UL	U	53	RoundHill	ALP	25/03/2009	L	ALP, TL, BTL	FALSE
25/03/2009	22/02/2009	108	UL	U	75	RoundHill	ALP	25/03/2009	L	ALP, TL, BTL	FALSE
25/03/2009	01/03/2009	86	UL	U	ARR	RoundHill	ALP	25/03/2009	L	ALP, TL	FALSE
25/03/2009	01/03/2009	86	UL	U	ARR	RoundHill	ALP	25/03/2009	L	ALP, TL	FALSE
27/03/2009	27/01/2009	92.5	UL	U	50	South Run	TL	27/03/2009	U	TL, BTL	TRUE
27/03/2009	27/01/2009	92.5	L	L	47	South Run	TL	27/03/2009	U	TL, BTL	FALSE
28/03/2009	01/03/2009	107.5	UL	U	95	Bonney Mo	ALP	28/03/2009	L	ALP, TL	FALSE
28/03/2009	01/03/2009	107.5	UL	U	91	Bonney Mo	ALP	28/03/2009	L	ALP, TL	FALSE
06/04/2009	27/01/2009	84	UL	U	78	South Run	TL	07/04/2009	U	TL, BTL	TRUE
06/04/2009	27/01/2009	84	UL	U	82	South Run	TL	07/04/2009	U	TL, BTL	TRUE
09/04/2009	27/01/2009	100	L	L	48	Abbott	ALP	09/04/2009	U	TL, BTL	FALSE
13/01/2008	Dec 5	131	L	E	45	Fidelity	TL	13/01/2008	E	TL, BTL	TRUE
13/01/2008	Dec 5	131	L	E	49	Fidelity	TL	13/01/2008	E	TL, BTL	TRUE
30/01/2008	Dec 5	72	UL	U	50	40 Watt	BTL	30/01/2008	U	TL, BTL	TRUE
30/01/2008	Dec 5	72	UL	U	51	40 Watt	BTL	30/01/2008	U	TL, BTL	TRUE
06/02/2008	Jan 26 SH	79	UL	E	ARR(45)	Fidelity	TL	06/02/2008	E	ALP, TL	TRUE
06/02/2008	Jan 26 SH	79	UL	E	ARR(33)	Fidelity	TL	06/02/2008	E	ALP, TL	TRUE

Continued...

Test Date	PWL	Layer Depth	PST	INT	% Cut	Site	Veg. Zone	Date	Prop	Applies to	MATCH
07/02/2008	Jan 26 SH	73	UL	E	ARR(38)	Camp West	TL	07/02/2008	E	ALP, TL	TRUE
07/02/2008	Jan 26 SH	73	UL	E	ARR(44)	Camp West	TL	07/02/2008	E	ALP, TL	TRUE
09/02/2008	Jan 26 SH	95-130	L	E	43	Fidelity	TL	09/02/2008	E	ALP, TL	TRUE
09/02/2008	Jan 26 SH	95-130	L	E	41	Fidelity	TL	09/02/2008	E	ALP, TL	TRUE
10/02/2008	Jan 26 SH	80-90	ARR	E	35% (47)	TCH cut	BTL	10/02/2008	E	ALP, TL	TRUE
10/02/2008	Jan 26 SH	80-90	ARR	E	42% (46)	TCH cut	BTL	10/02/2008	E	ALP, TL	TRUE
10/02/2008	Jan 26 SH	80-90	ARR	E	41% (46)	TCH cut	BTL	10/02/2008	E	ALP, TL	TRUE
10/02/2008	Jan 26 SH	80-90	ARR	U	46% (14)	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
10/02/2008	Jan 26 SH	80-90	ARR	U	46% (22)	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
10/02/2008	Jan 26 SH	80-90	L	L	37	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
10/02/2008	Jan 26 SH	80-90	L	E	46	TCH cut	BTL	10/02/2008	E	ALP, TL	TRUE
10/02/2008	Jan 26 SH	80-90	UL	U	92	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
10/02/2008	Jan 26 SH	80-90	UL	U	99	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
10/02/2008	Jan 26 SH	80-90	UL	U	87	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
10/02/2008	Jan 26 SH	80-90	UL	U	89	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
10/02/2008	Jan 26 SH	80-90	UL	U	86	TCH cut	BTL	10/02/2008	E	ALP, TL	FALSE
11/02/2008	Jan 26 SH	83	L	L	33	Illecillewaet	BTL	11/02/2008	E	ALP, TL	FALSE
11/02/2008	Jan 26 SH	83	L	E	48	Illecillewaet	BTL	11/02/2008	E	ALP, TL	TRUE
12/02/2008	Jan 26 SH	117	L	E	47	Fidelity	TL	12/02/2008	E	ALP, TL	TRUE
12/02/2008	Jan 26 SH	117	L	E	48	Fidelity	TL	12/02/2008	E	ALP, TL	TRUE
12/02/2008	Dec 5	225	L	E	41	Fidelity	TL	12/02/2008	U	TL, BTL	FALSE
13/02/2008	Jan 26 SH	140	L	L	37	South Run	TL	13/02/2008	E	ALP, TL	FALSE
13/02/2008	Jan 26 SH	140	UL	E	56	South Run	TL	13/02/2008	E	ALP, TL	TRUE
14/02/2008	Dec 5	218	L	L	38	Fidelity	TL	14/02/2008	U	TL, BTL	FALSE

Continued...

Test Date	PWL	Layer Depth	PST	INT	% Cut	Site	Veg. Zone	Date	Prop	Applies to	MATCH
14/02/2008	Jan 26 SH	127	L	L	25	Fidelity	TL	14/02/2008	E	ALP, TL	FALSE
16/02/2008	Jan 26 SH	108-118	UL	E	53	Abbott	ALP	15/02/2008	E	ALP, TL	TRUE
16/02/2008	Jan 26 SH	108-118	L	L	38	Abbott	ALP	15/02/2008	E	ALP, TL	FALSE
27/02/2008	Dec 5	250	SF	U	55 SF	Fidelity	TL	28/02/2008	U	TL, BTL	TRUE
18/03/2008	Dec 5	214	UL	U	51	Fidelity	TL	18/03/2008	U	TL, BTL	TRUE
18/03/2008	Feb 23 SH	80	L	L	44	Fidelity	TL	18/03/2008	L	ALP, TL	TRUE
19/03/2008	Feb 23 SH	90	L	L	50	Schuss Lk	TL	19/03/2008	L	ALP, TL	TRUE
19/03/2008	Feb 23 SH	90	UL	E	53	Schuss Lk	TL	19/03/2008	L	ALP, TL	FALSE
21/03/2008	Feb 23 SH	80	L	L	47	Roundhill	ALP	21/03/2008	L	ALP, TL	TRUE
24/03/2008	Feb 23 SH	98	L	L	48	Bostock Hd	TL	24/03/2008	L	ALP, TL	TRUE
24/03/2008	Feb 23 SH	98	L	L	46	Bostock Hd	TL	24/03/2008	L	ALP, TL	TRUE
26/03/2008	Feb 23 SH	77	L	L	50	Fidelity	TL	26/03/2008	L	ALP, TL	TRUE
26/03/2008	Feb 23 SH	77	UL	U	53	Fidelity	TL	26/03/2008	L	ALP, TL	FALSE
29/03/2008	Dec 5	231	UL	U	51	Fidelity	TL	29/03/2008	U	TL, BTL	TRUE
31/03/2008	Feb 23 SH	90	L	L	46	Bostock Hd	TL	31/03/2008	L	ALP, TL	TRUE
31/03/2008	Feb 23 SH	90	L	L	50	Bostock Hd	TL	31/03/2008	L	ALP, TL	TRUE