uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Marian Kremer

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.Sc. (Geography)

GRADE / DEGREE

Department of Geography

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Incorporation of Vegetation into Mountain Permafrost Distribution Models Southern Yukon Territory

TITRE DE LA THÈSE / TITLE OF THESIS

A. Lewkowicz

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

M. Sawada

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

K. Gajewski

S. Smith

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

Incorporation of Vegetation into Mountain Permafrost Distribution Models, Southern Yukon Territory

Marian Kremer

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
In partial fulfillment of the requirements
For the MSc degree in Geography

Department of Geography
University of Ottawa
Ottawa, Ontario, Canada
December, 2009

Your file Votre référence
ISBN: 978-0-494-66244-1
Our file Notre référence
ISBN: 978-0-494-66244-1

ABSTRACT

Three groups of variables (Digital Elevation Model [DEM]-derived variables, fieldwork-derived vegetation variables, and satellite imagery-derived vegetation variables) were combined in Classification and Regression Tree (CART) models to determine the utility of vegetation-based variables for mountain permafrost distribution modelling in the southern half of theYukon Territory. Four variables were measured in the field: canopy openness, vegetation height, organic mat thickness, and dominant species. Using Landsat TM and ETM+ imagery, three variables were calculated: a Normalized Difference Vegetation Index (NDVI), a vegetation classification, and a canopy closure classification. Individual variables were also examined to determine the one most useful for representing vegetation when modelling permafrost presence or absence. Additionally, models for each of five study areas spread across 5° of latitude were compared to examine the transferability of each variable.

The addition of vegetation variables to the CART models created with DEM-derived variables resulted in only a minimal increase in the overall accuracy. Dominant species proved to be the most useful variable, but the relationship between permafrost and each species differed among study areas. Only black spruce (*Picea mariana*) was consistently classified as permafrost probable and lodgepole pine (*Pinus contorta* var. *latifolia*) and trembling aspen (*Populus tremuloides*) were classified as permafrost improbable over all study areas. These results indicate that models of permafrost distribution across large areas are not likely to be enriched sufficiently by the inclusion of vegetation variables while models covering smaller areas may benefit from the inclusion of vegetation variables. The CART models tended to show a high accuracy in the prediction of areas with no permafrost which could be useful for the purposes of infrastructure development. CART models have not previously been used in permafrost modelling and the high accuracies they produced may indicate their utility for modelling the complex relationships among the variables affecting permafrost.

ACKNOWLEDGEMENTS

I wish to thank first, Dr. Antoni Lewkowicz and Dr. Mike Sawada, my supervisors, whose enthusiasm in this project inspired my interest. I gratefully acknowledge the financial support of the Canadian Foundation for Climate and Atmospheric Sciences (CFCAS), the Natural Sciences and Engineering Research Council of Canada (NSERC), my NSERC scholarship, the Federal Government International Polar Year (IPY) program, the Northern Scientific Training Program (NSTP), and the Faculty of Arts and the Faculty of Graduate and Postdoctoral Studies at the University of Ottawa, without which I could not have completed this project.

I would also like to thank Philip Bonnaventure, Andrea Sidler, and Jason Skucas for all their work and help in the field, and for the shared laughs (and silences) that made for such a memorable experience. Beyond the fieldwork, Phil has provided much advice, help, and support.

Finally I thank my parents for enduring the process, Sarah Moffat for providing motivation when I most needed it, Craig Smith for all the support, and the Friends who were there throughout my writing.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The relationship between vegetation and permafrost is complex, as each both affects and is affected by the other. Vegetation can promote the development, maintenance, or degradation of permafrost depending on the type or cover of the vegetation. Conversely, the conditions within the ground can affect the type and structure of vegetation above.

While elevation and Potential Incoming Solar Radiation (PISR) are widely used in mountain permafrost distribution models, vegetation is one variable that has not been examined independently or in depth. The inclusion of vegetation in models has been attempted on several occasions; however, these studies did not include detailed examinations of the interactions between permafrost and vegetation (Etzelmüller *et al.*, 2006; Gruber and Hoelzle, 2001; Leverington and Duguay, 1996; Morrissey *et al.*, 1986; Ødegård *et al.*, 1999). The variable has been omitted in many studies based on its apparent insignificant contribution, and in other cases it has been suggested as of potential value although not particularly significant to the study (Brenning *et al.*, 2005; Gruber and Hoelzle, 2001; Ødegård *et al.*, 1999).

The literature that discusses in detail the interaction between vegetation and permafrost does not focus on mountain permafrost modelling and was mainly written 20 to 50 years ago. This literature therefore appeared prior to the more recent advances in geographic information systems (GIS) and remote sensing technology that could perhaps help produce more detailed understanding of the relationship between the two variables. This research constitutes part of a project using GIS to develop a map of mountain permafrost within the southern Yukon. The goal for the larger project is to develop detailed estimates (30 x 30 m grid cells) of the probability of permafrost for an area of approximately 250 000 km^2. Currently the data used in the model to predict permafrost probability include basal temperature of snow (BTS) measurements (made in late-winter as an indicator of the presence of permafrost) and a Digital Elevation Model (DEM) used to provide information on elevation and PISR. Summer ground-truthing is used to develop a logistic regression model in which the probability of permafrost is related to the independent variables.

This section of the research examines the incorporation of vegetation data, both collected in the field and from satellite imagery, into mountain permafrost distribution models using Classification and Regression Trees (CART). Remotely sensed imagery, if used with an understanding of the relationships between vegetation and permafrost, may provide data that allow for the effective incorporation of continuous coverage of vegetation into a mountain permafrost distribution model.

1.1 Objectives

The goals of this research are to test the value of including vegetation variables, remotely sensed and collected in the field, in mountain permafrost distribution models and to determine the most useful vegetation variables for permafrost modelling in the southern half of the Yukon Territory. An additional component of the research is to determine whether the models and variables provide similar results across the different climatological regions of the study area or whether local conditions must be considered.

2 BACKGROUND

The distribution of permafrost that exists at elevations below the treeline cannot be modelled accurately solely using variables (e.g. elevation) that assume uniform surface conditions. Prediction of permafrost for areas where forest or shrubs dominate may be improved by accounting for the ground cover (Beilman and Robinson, 2003; Bliss, 1978; Brown, 1973; Leverington and Duguay, 1996; Morrissey *et al.*, 1986; Shur and Jorgenson, 2007; Viereck, 1973). In the southern half of the Yukon, vegetation may also be useful as an indicator of the subsurface conditions (Lewkowicz and Bonnaventure, 2008; Lewkowicz and Ednie, 2004).

The various methods for representing vegetation as a variable in permafrost models in previous studies and the differing results obtained provide an inconclusive understanding of how vegetation may best be fitted into a model of permafrost distribution in mountainous areas. Examining the interactions between permafrost and vegetation and the ways in which vegetation has been used in the past gives direction to the ways in which vegetation can potentially be used in permafrost distribution models.

2.1 Effects of Vegetation on Permafrost

Terrain factors including slope aspect and angle, surface roughness length, and soil properties (e.g., moisture content), as well as climatic conditions including temperature, cloud cover, prevailing winds, and precipitation, control the energy budget and the ways in which energy is exchanged within the environment (Oke, 1987; Steyn *et al.*, 1997). The presence of vegetation affects the partitioning of this balance. Vegetation has an impact on all fluxes in the energy balance equation including net radiation, sensible heat, latent heat, and ground heat flux and hence affects the ground thermal conditions.

Snow cover is also a link, albeit indirect, between vegetation and permafrost. Although snow cover is closely related to several parts of the surface energy balance equation, the relationship between vegetation, snow cover, and, ultimately, permafrost is complex and requires separate examination.

2.1.1 Net Radiation (Q)*

Vegetation affects net radiation in contradictory ways, as shown in Figure 2.1. The canopy shades the surface, reducing the heating of the ground by solar radiation (Brown, 1966; Brown and Péwé, 1973; Peddle and Franklin, 1993; Smith and Riseborough, 2002; Viereck, 1973; Yi *et al.*, 2007). Cooler ground surface temperatures inhibit thawing in shaded areas, as compared to areas where the surface is free of vegetation (Brown, 1966; Brown and Péwé, 1973; Heggem *et al.*, 2006; Viereck, 1973; Yi *et al.*, 2007). The effects of shading on ground temperatures are more significant during the summer, when incoming solar radiation is contributing to thawing and warming the ground (Brown, 1966; Rouse, 1982; Viereck, 1973). Shading is of lesser importance in the winter because other factors, such as thickness of snow cover and ground thermal conductivity, play a more prominent role in controlling ground temperatures (Brown, 1966). Also, in areas with broadleaf vegetation, canopy cover is reduced and therefore of less significance in winter (Brown, 1966).

Brown (1966) also indicates that albedo of the ground cover can affect net radiation, but only minimally. Lighter coloured plants, such as lichens, reflect more radiation and allow less to be absorbed into the ground (Brown, 1966). The result would be cooler ground temperatures and possible development or maintenance of permafrost.

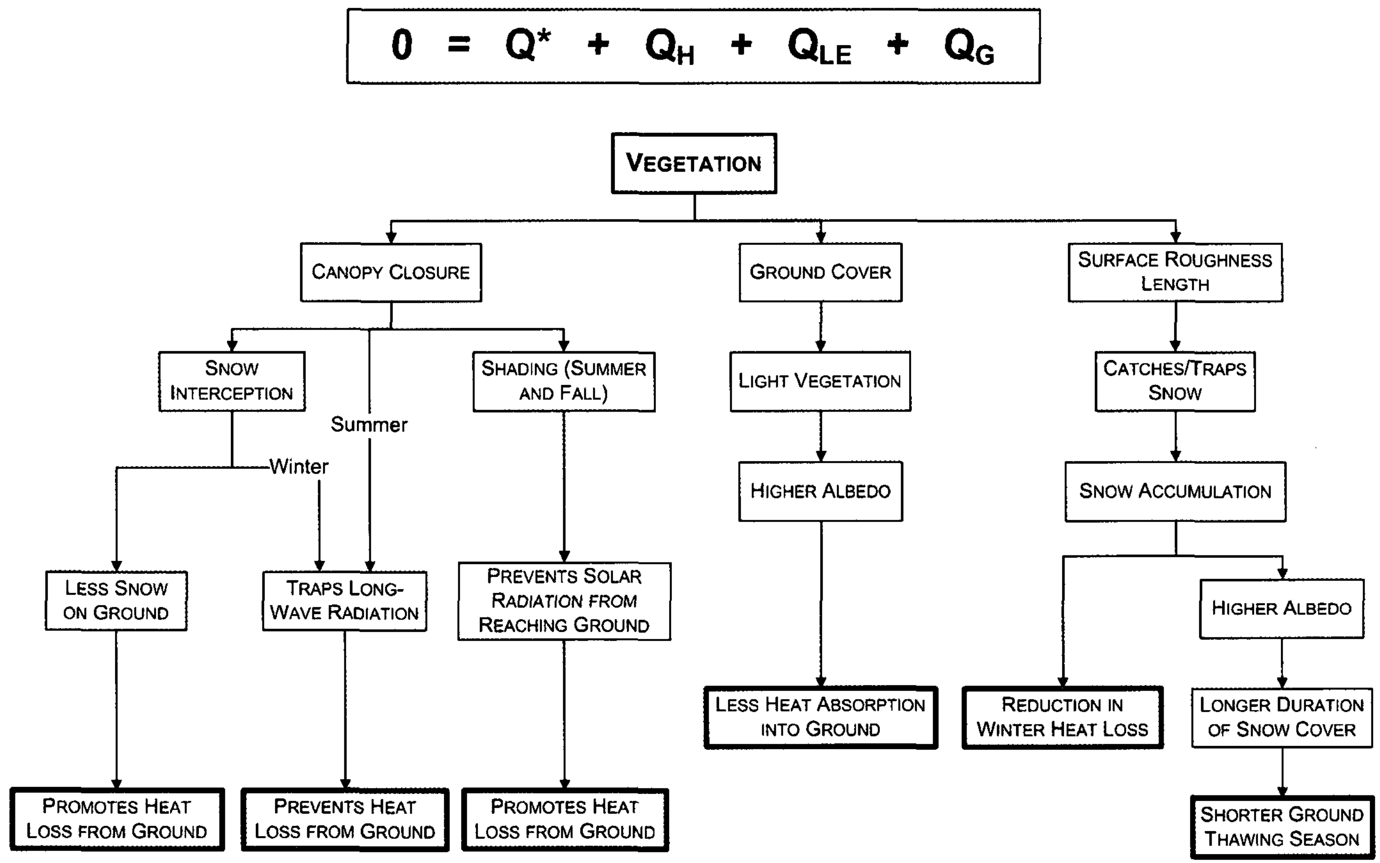


Figure 2.1: Effects of vegetation on net radiation and the consequent impacts on permafrost.
Blue boxes indicate parts of the system that promote permafrost development or maintenance while red boxes promote the maintenance of seasonally frozen ground or permafrost degradation.

2.1.2 Turbulent Fluxes (Q_H and Q_{LE})

Vegetation is an influential factor with regard to the sensible (Q_H) and latent heat (Q_{LE}) exchange portions of the energy balance equation, as shown in Figures 2.2 and 2.3, respectively. Of primary importance here is the surface roughness length created by vegetation, which affects wind speed at the ground surface (Lee, 1978; Oke, 1987). Loss of heat from the ground surface, through free convection, is not particularly efficient; however, wind turbulence above the boundary layer carries away the warm air, increasing the efficiency of the process (Lee, 1978).

In a vegetated environment like a forest, wind speeds below the canopy tend to be lower than above because the vegetation creates a thick boundary layer, approximately two thirds the height of the forest stand (Brown, 1966; Brown and Péwé, 1973; Harding *et al.*, 2002; Oke, 1987). Sensible heat must be inefficiently convected through this layer before its transfer can be assisted by the turbulent air above (Lee, 1978).

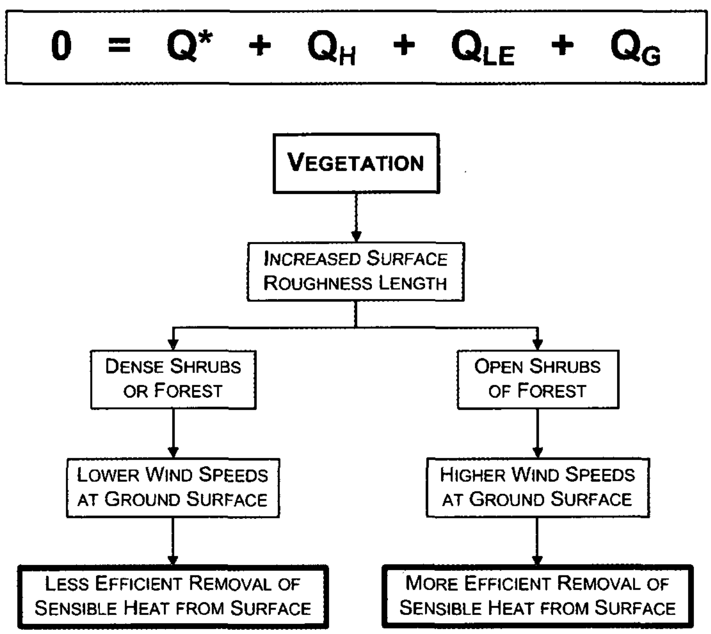


Figure 2.2: Effects of vegetation on sensible heat exchange and the consequent impacts on permafrost.
Blue boxes indicate parts of the system that promote permafrost development or maintenance while red boxes promote the maintenance of seasonally frozen ground or permafrost degradation.

Consequently, there is less efficient removal of sensible heat from the ground surface under a forest than on bare ground. The ground in a forest stand will therefore tend to stay warmer in the summer, creating conditions that are detrimental to the development or maintenance of permafrost. In open forest stands, however, the wind can pass between the trees at a much lower level (Brown, 1966, 1970). This results in more efficient removal of heat from the ground surface and greater potential for permafrost development or maintenance. Similarly, over an unvegetated surface, where the boundary layer is thinner, higher wind speeds can be found closer to the ground (Lee, 1978; Oke, 1987). At these sites, because heat is convected through a very thin layer of unmoving air, the sensible heat exchange can be a relatively efficient process.

Latent heat exchange is affected by vegetation in a similar manner to sensible heat exchange (Figure 2.3). In this case, however, heat transfer occurs as a result of the vapour pressure gradient between the moist surface and the relatively dry air above (Geiger *et al.*, 2003; Lee, 1978; Oke, 1987). Above the boundary layer, air turbulence

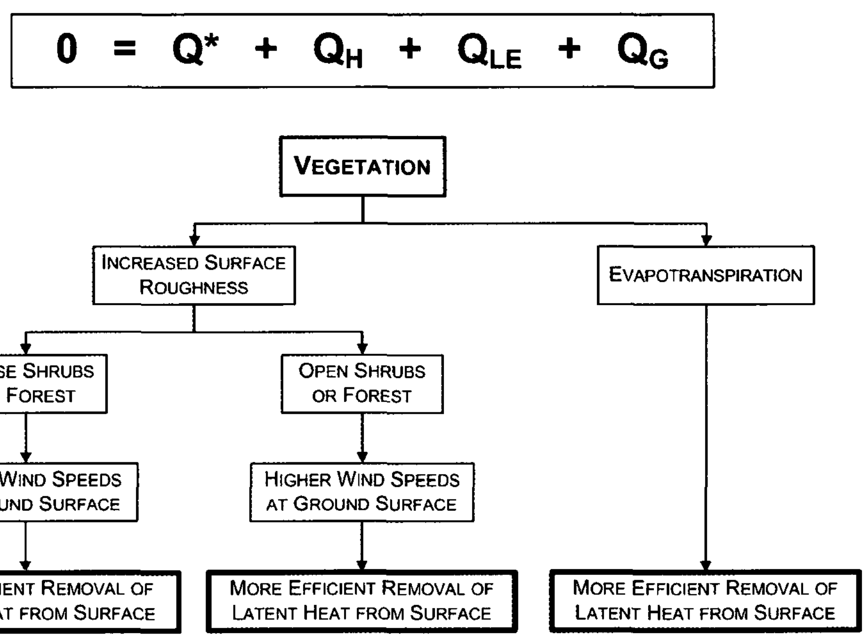


Figure 2.3: Effects of vegetation on latent heat exchange and the consequent impacts on permafrost. *Blue boxes indicate parts of the system that promote permafrost development or maintenance while red boxes promote the maintenance of seasonally frozen ground or permafrost degradation.*

efficiently carries away the moist air and replaces it with relatively drier air, allowing more evaporation to occur within the boundary layer (Geiger *et al.*, 2003; Lee, 1978; Oke, 1987).

Evaporation is less efficient under a forest than on bare ground. Moisture from the ground must travel through a thick boundary layer before it reaches the turbulent air above (Lee, 1978; Oke, 1987). The ground surface loses less heat, but retains its moisture longer, which limits the amount of energy being used to warm the ground (Rosenberg *et al.*, 1983). Within a forest, other mechanisms of moisture removal are more efficient than surface evaporation. Transpiration carries moisture from the ground through the roots and stems and out to the leaves of the vegetation, where the moisture is then evaporated (Brown, 1966; Lee, 1978; Oke, 1987). Because evaporation from the leaves is occurring at or closer to the active surface layer, this is a more efficient process of removing heat from the ground (Geiger *et al.*, 2003; Lee, 1978).

Over unvegetated surfaces, moisture is carried away by the wind more efficiently and the thin boundary layer is replenished with dry air at a faster rate (Lee, 1978; Oke, 1987). The ground, however, may also dry out at a faster rate, subsequently allowing excess radiant energy to be used to warm the ground.

Sensible and latent heat exchange help reduce the amount of energy that goes into warming the ground, thus maintaining a cooler ground surface temperature. As long as vegetation conditions allow for relatively efficient transfer of heat from the ground, the ground temperature will remain cooler, facilitating the potential development of permafrost.

2.1.3 Ground Conditions (Q_G)

The ground thermal characteristics below the surface also have an impact on the energy balance (Figure 2.4). Vegetation, including any moss layer on the ground surface, supports development of permafrost and thinning of the active layer (Bliss, 1978; Brown, 1966; Brown and Péwé, 1973; Cowell *et al.*, 1978; Rouse, 1982; Shur and Jorgenson, 2007; Viereck, 1973, 1982; Yi, *et al.*, 2007). Because peat is closely related to vegetation cover in its properties and effects on the ground thermal regime, it is included in this discussion.

A thicker organic mat is often associated with permafrost presence as it usually provides more insulation for the ground below (Beilman and Robinson, 2003; Bliss, 1978; Brown, 1966, 1973, 1978; Cowell *et al.*, 1978; Viereck, 1973, 1982). The thickness of the organic mat can be associated with the vegetation: shrub and tundra vegetation tend to have the thinnest organic layers, broadleaf forests tend to

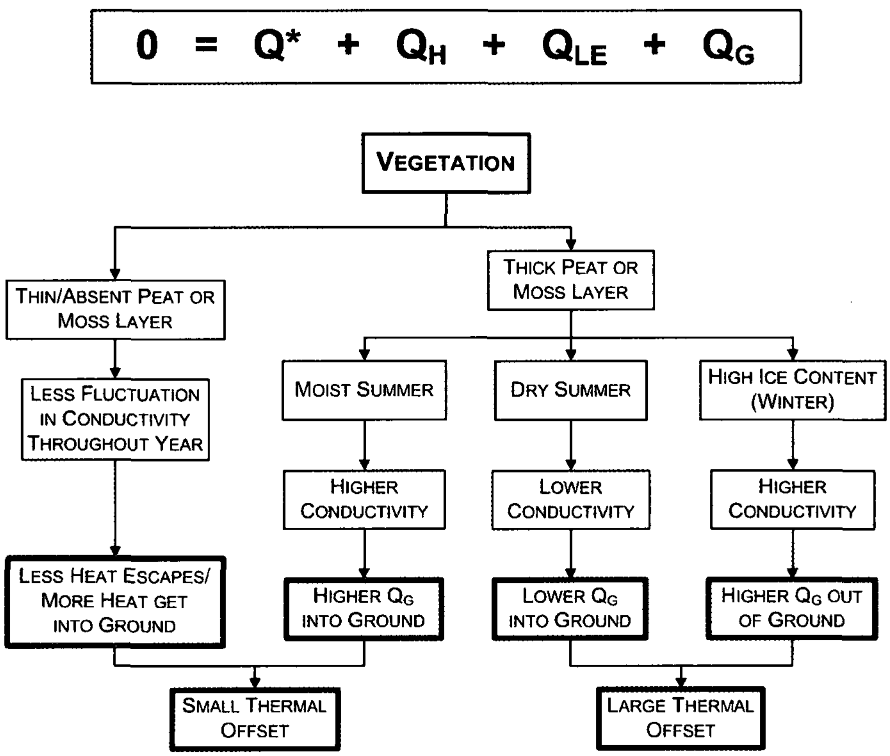


Figure 2.4: Effects of vegetation on ground heat exchange and the consequent impacts on permafrost. *Blue boxes indicate parts of the system that promote permafrost development or maintenance while red boxes promote the maintenance of seasonally frozen ground or permafrost degradation.*

have a thicker organic layer, and needle-leaf forests usually have the thickest layers (Brown, 1966, 1973).

The surface vegetation, organic material, and peat affect the thermal conductivity of the ground, which in turn controls the exchange of heat between the air above and the ground below the surface (Bliss, 1978; Brown, 1966; Brown and Péwé, 1973; Cowell *et al.*, 1978; Rouse, 1982; Viereck, 1973, 1982). This material can hold large quantities of water, which, when frozen, increases the thermal conductivity of the ground, facilitating the extraction of heat from the ground during the winter (Bliss, 1978; Brown, 1966; Brown and Péwé, 1973; Carey and Woo, 2001; Rouse, 1982; Smith and Riseborough, 2002; Yi *et al.*, 2007). In summer, however, the surface material may be drier due to evaporation of surface moisture (Brown, 1966; Brown and Péwé, 1973). Dry peat is a very poor conductor, so less heat can penetrate into the ground during the summer (Brown, 1966; Smith and Riseborough, 2002; Yi *et al.*, 2007).

The conductivity of the mineral soil can also affect the presence or absence of permafrost. As with peat or moss layers, moisture content is a major influence on conductivity, and additional effects are produced by the mineral composition. Texture affects how much water the soil can hold (Lee, 1978; Rosenberg *et al.*, 1983). Coarse-textured soils like sand have less intergranular contact and, when dry, provide greater insulation and hence lower conductivity. In contrast, fine grained soils like silt and clay have higher conductivities (Rosenberg *et al.*, 1983). The addition of moisture to these soils increases conductivity significantly (Rosenberg *et al.*, 1983). The ability of the soil to hold moisture is of importance in this regard. Coarser-grained materials can hold less moisture than finer material, and they tend to drain more efficiently. Thus, coarser-grained materials tend to be less conductive in the field (Rosenberg *et al.*, 1983). However, the coarse-grained, dry soils heat up more in the summer than fine-grained, moist soils because there is less moisture to evaporate.

2.1.4 Snow Cover

Apart from its direct influence, vegetation also indirectly affects net radiation through its influence on snow cover (Figure 2.1). Two contrasting observations about the influence of vegetation on snow depth have been described in the literature: the canopy acting as a shield, blocking accumulation of snow on the ground surface; and vegetation acting as a net, trapping and increasing surface accumulation of snow. The depth of the snow in a vegetated area is related to the species and, consequently, the structure or density of the vegetation present (Clark *et al.*, 1985; Pomeroy and Goodison, 1997).

During winter, a dense vegetation canopy can intercept the snowfall, reducing accumulation on the ground (Brown, 1966, 1970; Geiger *et al.*, 2003; Harding *et al.*, 2002; Oke, 1987; Peddle and Franklin, 1993; Pomeroy and Goodison, 1997; Shur and Jorgenson, 2007; Woo and Marsh, 2005; Yi *et al.*, 2007). This interception is particularly characteristic of needle-leaf forests (Pomeroy and Goodison, 1997; Shur and Jorgenson, 2007). The thinner layer of snow provides less insulation, allowing heat to be removed from the ground more efficiently than in areas of deeper snow accumulation (Brown, 1966; Heggem *et al.*, 2006; Yi *et al.*, 2007; Zhang *et al.*, 1997). Snow trapped in the canopy, however, helps to act as a barrier to long-wave radiation, reducing its escape to the atmosphere (Brown, 1966; Harding *et al.*, 2002; Rosenberg *et al.*, 1983). This means that the ground loses heat less efficiently and may actually stay warmer than in more open areas.

In contrast, slightly more open forest or shrub cover can catch and trap drifting snow (Clark *et al.*, 1985; Gray and Brown, 1982). Wind moves snow from the smoother, more exposed landscape (Pomeroy *et al.*, 2006; Zhang *et al.*, 1997). When the snow reaches the denser, taller vegetation, the increased surface roughness causes the wind speed to decrease sufficiently for deposition to occur (Brown, 1966; Pomeroy *et al.*, 2006; Pomeroy and Goodison, 1997; Woo and Marsh, 2005). As well, the tangle of branches traps snow and prevents it from being blown away (Brown, 1966;

Zhang *et al.*, 1997). The result is a deeper layer of snow, which insulates the ground from cold air temperatures and prevents heat in the ground from being transferred to the air (Gray and Brown, 1982; Harding *et al.*, 2002; Heggem *et al.*, 2006; Rouse, 1982; Smith and Riseborough, 2002; Viereck, 1973; Yi *et al.*, 2007; Zhang *et al.*, 1997). More of the heat absorbed during the thawing season is retained within the ground, keeping it warmer than uninsulated or less insulated surfaces (Brown, 1966). Heat cannot be released as easily during the winter under deep snow cover, so the only time during which there is efficient heat transfer is in summer, when heat is being transferred into the ground. The net effect of thick winter snow cover in lowland areas is that permafrost is less likely (Brown, 1966; Gray and Brown, 1982; Leverington and Duguay, 1996).

While a thick snow cover insulates the ground and warms it, a very thick accumulation may have the opposite effect by shortening the length of the thawing season since ground thaw cannot occur before the snow cover has ablated. In mountain areas, thick accumulations of snow may not melt until late summer. The greater snow accumulation has a high albedo, reflecting the majority of incoming solar radiation and thus reducing the energy absorbed (Brown, 1966, 1970; Harding *et al.*, 2002). In forest or high shrubs, the snow melt is also inhibited by shading from the canopy (Geiger *et al.*, 2003), and the ground thawing season is consequently shorter (Brown, 1966; Gray and Brown, 1982; Rouse, 1982).

2.2 Effects of Permafrost on Vegetation

Vegetation can affect permafrost presence or absence, but permafrost can also have an impact on vegetation (Figure 2.5). Permafrost, because of the associated cold ground temperatures and inefficient drainage, limits the size and number of species of plants that can grow above it (Brown, 1970; Brown and Péwé, 1973). Roots commonly are shallow because of the presence of impenetrable frozen ground and cold soil temperatures (Tryon and Chapin, 1983). Growth is limited to the horizontal plane, which does not provide a sufficiently stable base for many larger species of plants (Brown and Péwé, 1973; Tryon and Chapin, 1983). This explains the typical absence of taller forest stands in areas underlain by permafrost and the "drunken" structure that is sometimes associated with an insufficient root base (Brown, 1970; Brown and Péwé, 1973; Kokelj and Burn, 2003). In general, forest stands over permafrost with a thin active layer tend to be more open and their growth is stunted (Bliss, 1978; Brown, 1970; Kokelj and Burn, 2003).

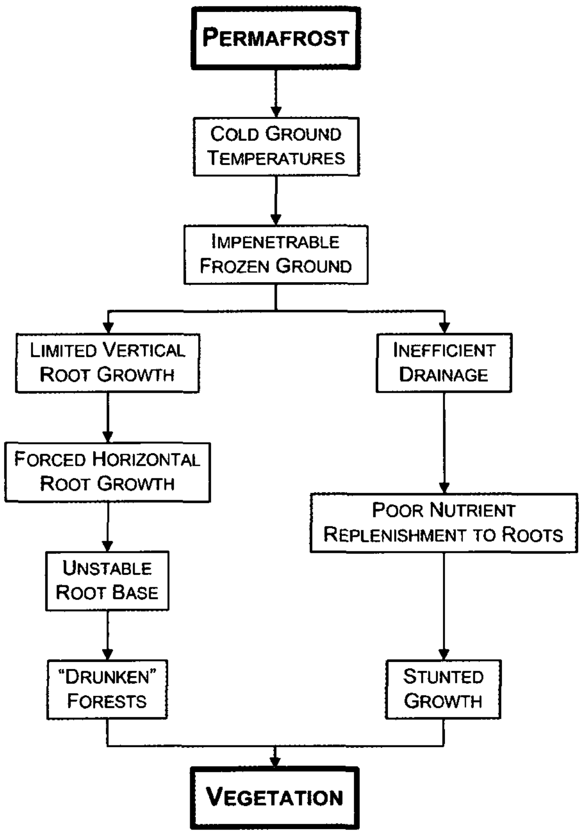


Figure 2.5: Effects of permafrost on vegetation.

Each plant species has preferred growing conditions. In general, only those that can tolerate the harsh ground conditions produced by permafrost can grow in such terrain. Broadleaf trees tend to grow in warmer, well-drained soils where permafrost tends to be absent, or they require a thick active layer (Bliss, 1978; Brown, 1973; Leverington, 1995; Viereck, 1973). Needle-leaf species, particularly black spruce (*Picea mariana*), are more often associated with permafrost presence. Black spruce forests tend to be found in more poorly drained soils with a thin active layer (Beilman and Robinson, 2003; Bliss, 1978; Duchesne *et al.*, 2003; Leverington, 1995; Viereck, 1973). White spruce (*Picea glauca*), in contrast, is more often found in slightly warmer, better drained soils (Bliss, 1978). While it is still often associated with permafrost, the active layers may be deeper (Bliss, 1978; Viereck, 1973).

2.3 Local and Global Scale Variations

There are differences in the interactions between permafrost and vegetation at both local and global scales. Interestingly, however, some factors, including vegetation type and soil moisture content, do not appear to vary greatly around the Arctic.

Although in many cases PISR (predicted assuming a bare ground surface) can be of great value for modelling permafrost distribution, it does not always provide all of the information. For example, in the southwest Yukon, thinner active layers have been found on southeast-facing slopes that have dense vegetation (Brown and Péwé, 1973). In this case, PISR would have provided results suggesting that permafrost would likely be absent because the slopes receive considerable amounts of solar radiation; however, the dense vegetation cover shaded the surface enough to allow permafrost to develop. This is not consistent, however, with most findings in the Yukon. Leverington (1995) found that, in the area around Mayo, spruce forests underlain by permafrost tended to dominate on north-facing slopes, while broadleaf forests with no permafrost below were found on the south-facing slopes.

In North America, vegetation and permafrost overlap in many areas, including in the mountains. One exception to this trend was found by Gray and Brown (1982) in the Chic-Choc Mountains in eastern Canada. Here, the boundary of the permafrost roughly follows the edge of the continuous vegetation cover. This occurrence was explained by the large amount of snow that was blown away from the unvegetated areas and trapped by the forest. The result was a thicker accumulation of snow in the more continuously vegetated areas, more insulation in the winter, and, consequently, warmer, permafrost-free soils in the vegetated areas (Gray and Brown, 1982).

In the European Alps, the vast majority of permafrost is found above the treeline, similar to the pattern found in the Chic-Choc Mountains (Brenning *et al.*, 2005; Gruber and Hoelzle, 2001; Kneisel, 2003; Ødegård *et al.*, 1999). Only low shrubs and tundra vegetation can be found in the majority of the areas underlain by permafrost. This vegetation cover does not appear to have a particularly significant effect on the permafrost below, as it has had limited utility for predicting the distribution of alpine permafrost (Brenning *et al.*, 2005; Gruber and Hoelzle, 2001; Ødegård *et al.*, 1999). Although the strong correlation between vegetation, elevation, and permafrost prevents vegetation from having much influence in the Alps, patches of permafrost below the treeline do appear to be influenced by vegetation. Kneisel (2003) reports that very sporadic permafrost can appear below the treeline in well-shaded areas.

Bliss (1978) suggests similarities between the vegetation-permafrost interactions in North America and those in Siberia. In both areas, needle-leaf trees, although of different species, are associated with the presence of permafrost. The findings of Etzelmüller *et al.* (2006) show that the relationship between permafrost and vegetation in mountains in northern Mongolia is similar to that in North America. In this area, it was found that south-facing slopes were less likely to have permafrost, and the vegetation, mainly grasses and open forest, was indicative of drier soil. The forested north-facing slopes, in contrast, were the most likely to be underlain by permafrost (Etzelmüller *et al.*, 2006).

The main difference between the research described in the literature and this project. This project attempts to model permafrost at a finer resolution than many previous studies and across multiple, widely spaced study areas that correspond to different climatological regions in the Yukon. This type of analysis is essentially unique.

2.4 Disturbance to Vegetation

A number of studies have discussed the effects of vegetation disturbance on permafrost (Brown, 1966; Shur and Jorgenson, 2007; Viereck, 1973, 1982). Disturbances that may have a noticeable impact on permafrost conditions are those that either remove vegetation layers, such as fire or logging activities, or those that alter the structure or health of the vegetation, like floods or wind throw (Brown, 1966; Viereck, 1973, 1982). Warm water from flooding that penetrates into the ground also encourages thawing of the permafrost (Viereck, 1973). As well, sediment from flood water can build up on the moss layer that insulates the ground and kill it (Viereck, 1973). Fires destroy the layers of vegetation, and potentially remove the insulating ground surface cover as well as the canopy shading (Shur and Jorgenson, 2007; Viereck, 1973, 1982; Yi *et al.*, 2007). Removal of vegetation by humans has similar results: reduced insulation and protection of the ground below, reduced shading of the surface, and consequently a deeper active layer (Brown, 1966; Yi *et al.*, 2007). If left to natural processes, the affected area becomes re-vegetated through succession, and the insulating and protective properties of the returning vegetation can help promote thinning of the active layer (Shur and Jorgenson, 2007; Viereck, 1973).

As discussed by Shur and Jorgenson (2007), the five types of permafrost formation help demonstrate the permanency of the changes to permafrost that occur due to vegetation disturbance and subsequent successional stages. Figure 2.6 shows the figure by Shur and Jorgenson (2007) showing how each type of permafrost is associated with vegetation succession and climate conditions. "Climate-Driven" and "Climate-

Driven, Ecosystem-Modified" permafrost can recover as long as the climatic conditions are favourable for the development of permafrost (Shur and Jorgenson, 2007). In areas where climatic conditions are not (or no longer) conducive to the development of permafrost, "Climate-Driven, Ecosystem-Protected" and "Ecosystem-Driven" permafrost can recover as long as the vegetation conditions return to their previous state prior to the disturbance (Shur and Jorgenson, 2007). "Ecosystem-Protected" permafrost will be lost if a vegetation disturbance occurs (Shur and Jorgenson, 2007).

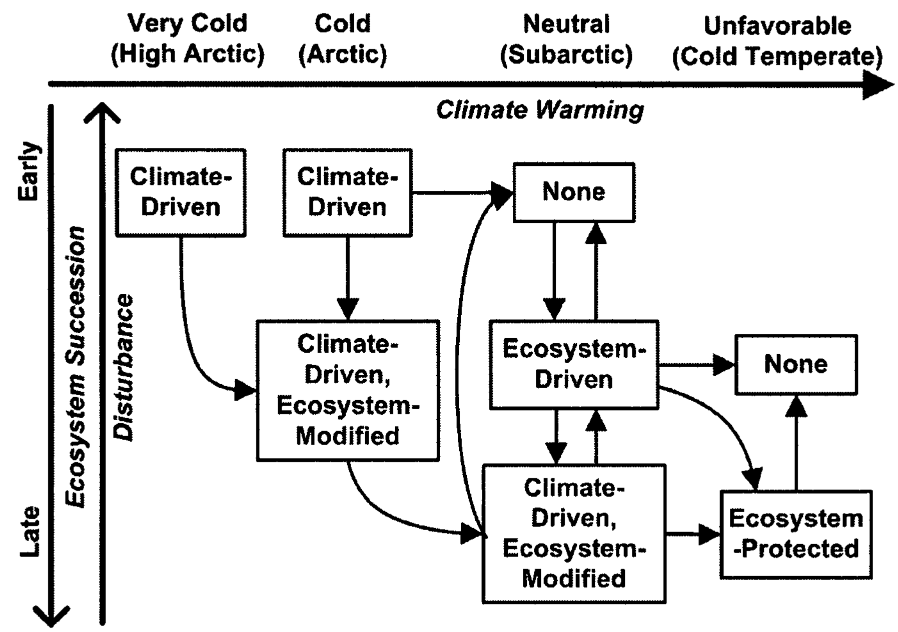


Figure 2.6: Permafrost conditions in relation to climate, ecological succession, and disturbance (From Shur and Jorgenson, 2007).

2.5 Previous Research on Vegetation and Permafrost Modelling

Attempts to include vegetation in mountain permafrost models have not been numerous, and no generally accepted method exists at this point. The studies that have included vegetation have tended to use remotely sensed vegetation data, either vegetation indices or landcover classifications. Vegetation indices such as the Normalized Difference Vegetation Index (NDVI) and the Soil Adjusted Vegetation Index (SAVI) have been tested in BTS models in Europe, but with little success. Ødegård *et al.* (1999) tested NDVI in their permafrost model in southern Norway, but it provided a statistically insignificant contribution to a model that already contained elevation and apparent satellite temperatures (extracted from the thermal band of Landsat imagery). This insignificance was likely due to the strong correlation between elevation and NDVI in the study area. Gruber and Hoelzle (2001) also found that the vegetation index (in this case, SAVI) was not of great value in a BTS permafrost model in Switzerland. The r^2 value of the model containing SAVI, elevation, and solar radiation was 0.398, which is only a small increase over the r^2 value of 0.386 achieved using only elevation and solar radiation. In fact, Brenning *et al.* (2005) discarded the vegetation variable because of its lack of significance in the original study by Gruber and Hoelzle (2001). In each of these cases, vegetation was strongly correlated with elevation, and thus provided little additional explanatory value to the model.

In contrast to the vegetation indices, landcover classifications, derived from satellite imagery, have shown promising results for permafrost distribution modelling. Models that include landcover classifications for this purpose, however, are limited in output, usually producing binary results (presence or absence of permafrost) that potentially limit their accuracy and value.

In the Yukon-Tanana Uplands of Alaska, Morrissey *et al.* (1986) divided landcover into nine classes of vegetation that included types of canopy cover (closed or open) and types of vegetation (needle-leaf forest, broadleaf forest, mixed forest, shrub) using Landsat imagery. They employed logistic discriminant functions using the landcover classification and data from the thermal band of Landsat imagery. The logistic discriminant functions allowed for the use of discrete datasets such as the landcover classification. The predicted permafrost distribution was reasonably successful. However, the output was limited to only three permafrost classes: frozen, discontinuously frozen, and unfrozen. In the Mayo region in the Yukon (a lowland area), Leverington and Duguay

(1996) also used a vegetation classification from Landsat imagery along with several other variables including NDVI, equivalent latitude (essentially PISR), and slope aspect. In this case, they were testing variables and three supervised classification methods, two of which could be used with discrete data, for predicting active layer depths and presence or absence of permafrost. The landcover classification was found to be the most useful predictor. The results showed that modelling two classes (frozen and unfrozen ground) had fairly high accuracy (up to 93%), but three depth classes proved to be considerably less accurate (56%- 70%). Etzelmüller *et al.* (2006) used a combination of variables, including terrain characteristics, a landcover classification, and NDVI, for their multi-criteria analysis of mountain permafrost distribution in Mongolia. Based on permafrost favourability values calculated using the multiple criteria, they were able to differentiate between four levels of permafrost presence: unlikely permafrost, possible permafrost, probable permafrost, and highly probable permafrost. They found that forest cover and density, factors that can be derived from a combination of the vegetation classification and NDVI, were particularly useful for determining permafrost distribution.

2.6 Vegetation as an Indicator of Permafrost

Based on the relationships between permafrost and vegetation, it appears that there are several ways in which remotely sensed vegetation data can be used as an indicator of permafrost presence or absence.

The canopy cover may be a valuable indicator of the distribution of permafrost. Canopy cover, as measured in terms of open or closed forest stands, relates to how much solar radiation can actually reach the ground surface. A closed forest will allow less radiation to reach the surface, thus allowing less heat to penetrate into the ground (Figure 2.1). An open canopy, however, allows higher wind speeds to reach the ground surface, which aids in the transfer of sensible and latent heat from the ground (Figures 2.2 and 2.3) (Brown, 1966). From below the surface, the cool ground temperatures and often poor drainage limit the height and density of tree growth above. Thus, an open forest stand may be an indicator of permafrost (Figure 2.5) (Bliss, 1978; Brown, 1970; Brown and Péwé, 1973; Kokelj and Burn, 2003; Tryon and Chapin, 1983).

Ground cover appears to be one of the most influential and noticeable factors in the relationship of vegetation to permafrost. The thickness of the peat or moss layer on the ground is of particular importance with regard to near-surface thermal conductivity. It acts as an insulator in summer, preventing heat from penetrating deep into the ground, and as a good conductor in winter, allowing heat to escape readily, resulting in a cooler ground temperature throughout the year (Brown, 1966, 1973; Brown and Pewe, 1973; Cowell *et al.*, 1978; Viereck, 1973). Although the organic mat thickness is of great importance with regard to the subsurface conditions, this variable can only be measured through extensive field work or through a proxy variable that can be mapped over a large area. Vegetation type may be of great value as an indicator of these conditions (Dingman and Koutz, 1974; Duchesne *et al.*, 2003; Leverington and Duguay, 1996). Different plant species prefer different types of soil conditions and produce different thicknesses of organic mats. For example, broadleaf trees tend to grow in areas with well-drained soils and no permafrost or a thick active layer, while needle-leaf species like black spruce are commonly found in poorly drained, wetter soils with a thin active layer (Beilman and Robinson, 2003; Bliss, 1978; Carey and Woo, 2001; Dingman and Koutz, 1974; Leverington and Duguay, 1996; Morrissey *et al.*, 1986; Shur and Jorgenson, 2007; Viereck, 1973). Permafrost is often overlain by a thick active layer or is absent in areas where the vegetation cover includes white spruce, trembling aspen (*Populus tremuloides*), or paper birch (*Betula papyrifera*) (Bliss, 1978; Brown, 1973; Leverington and Duguay, 1996; Viereck, 1973).

The differentiation between broadleaf and needle-leaf forest may also be useful with regard to snow depth. Clark *et al.* (1985) and Pomeroy and Goodison (1997) report that broadleaf forest appears to be more efficient in trapping and holding snow as compared to open areas and needle-leaf tree species. Dense needle-leaf forests are associated with interception of snow by the canopy and reduced snow accumulation on the ground (Brown, 1966, 1970; Geiger *et al.*, 2003; Oke, 1987; Pomeroy and Goodison, 1997). However, these findings are not consistent through the literature. Gray and Brown (1982) and Rouse (1982) found that the needle-leaf forests in their study areas trapped blowing snow and thus increased accumulation.

A vegetation classification that differentiates between broadleaf and needle-leaf vegetation as well as between open and closed stands may provide indications about ground conditions and ultimately the distribution of permafrost. Open canopy, needle-leaf forest is most likely to be underlain by permafrost (Beilman and Robinson, 2003; Bliss, 1978; Brown, 1966, 1973; Cowell *et al.*, 1978; Viereck, 1973). Classes for bare ground and shrub vegetation would also be useful as indicators of the potential snow cover and ground insulation throughout the winter.

As discussed above, healthier, more productive vegetation is often located in areas with either very deep active layers or no permafrost at all. Therefore, a vegetation index measuring the productivity of the vegetation might prove to be a useful indicator of the distribution of permafrost. One concern about this particular measure, however, is that confusion may occur between areas where the vegetation has been infected with disease or has been burned, and the areas underlain by permafrost. It is also a temporally sensitive variable due to autumn senescence and spring leaf-out. Imagery from approximately the same time of year, early August, with areas affected by disease and burns masked out, would help prevent errors caused by the above factors.

3 STUDY AREAS

Five field areas within the southern Yukon, identified for the main permafrost distribution modelling project, were examined for this research: Dawson, Faro, Keno, Johnson's Crossing, and the Sa Dena Hes mine site north of Watson Lake (Figure 3.1). These areas are widely spaced, covering four of the major climatological regions of the southern Yukon (Wahl *et al.*, 1987), and are located within the zones of sporadic and widespread discontinuous permafrost (Heginbottom *et al.*, 1995). The areas cover four different ecoregions within the Boreal Cordillera ecozone (Environment Canada, 2009a). These study areas are located within mountainous terrain, with elevations ranging between approximately 600 to 2000 m asl with the exception of the Dawson area, which extends down to approximately 270 m asl. Permafrost is present at higher elevations, but it is also found below the treeline due to cold air drainage and hydrological conditions.

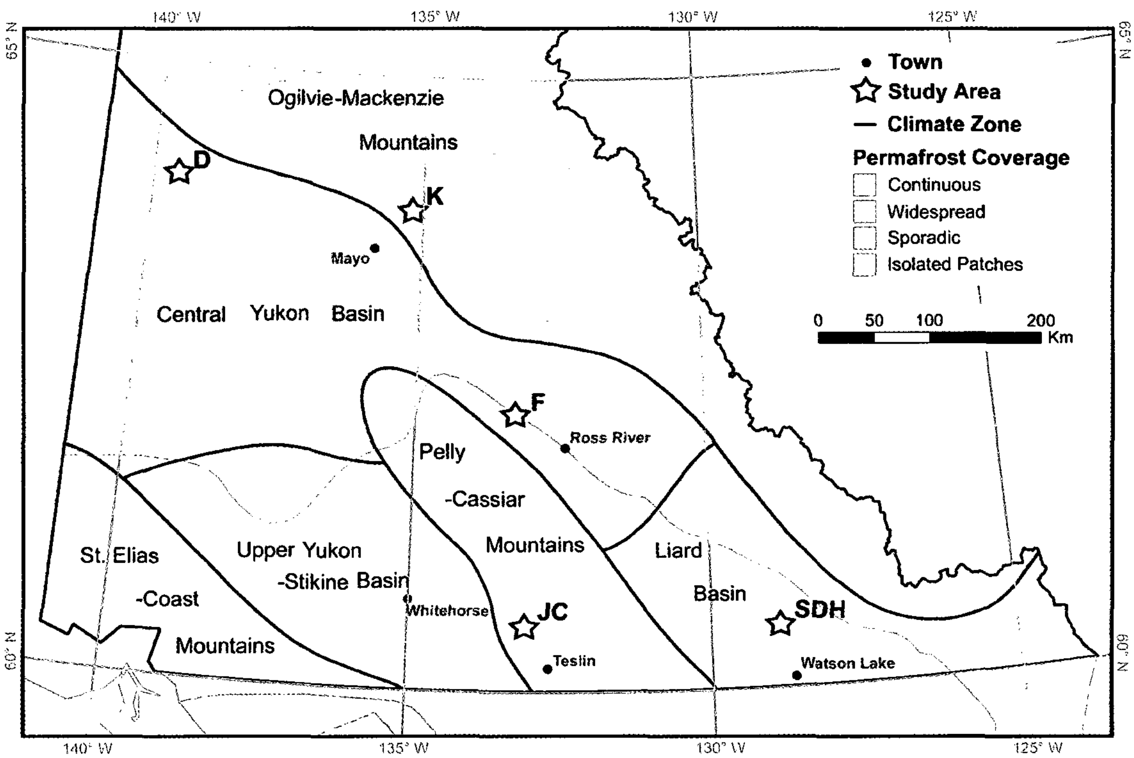


Figure 3.1: Location of each study area in relation to permafrost zones and climate regions (Permafrost zones: Heginbottom *et al.*, 1995; Climate Regions: Wahl *et al.*, 1987).
D: Dawson, F: Faro, K: Keno, JC: Johnson's Crossing, SDH; Sa Dena Hes

3.1 Dawson

The Dawson study area (Figure 3.2) is centred at 64° 7.3' N, 139° 41.8' W in the vicinity of the town site and along the Top-of-the-World Highway. The entire study area is approximately 1130 km^2 and ranges in elevation from approximately 270 to 1255 m asl. It is within the zone of widespread discontinuous permafrost (50-90% underlain by permafrost) (Heginbottom *et al.*, 1995). It is located in the north-west of the Central Yukon Basin climatological region (Figure 3.1), which is characterized by warm summers, cold winters, with moderate precipitation (300 - 400 mm per year) (Wahl *et al.*, 1987). The mean annual air temperature (MAAT) (from 1971 to 2000) recorded at the Dawson Airport weather station (64° 02' N, 139° 07' W, elevation 690.4 m asl) is -4.4°C, and the mean annual precipitation is 324 mm (Environment Canada, 2009b). Vegetation of the Klondike Plateau ecoregion typically includes black spruce, paper birch, white spruce, trembling aspen, and some lodgepole pine (*Pinus contorta* var. *latifolia*), as well as scrub birch (*Betula glandulosa* var. *glandulosa*) and willow (*Salix* spp.) (Environment Canada, 2009a). Few areas with tundra vegetation were found during fieldwork within this study area: low scrub birch and willow were dominant above treeline. This study area differs from all the others because it was never glaciated, and hence is characterized by V-shaped valleys and rolling hills (Environment Canada, 2009a).

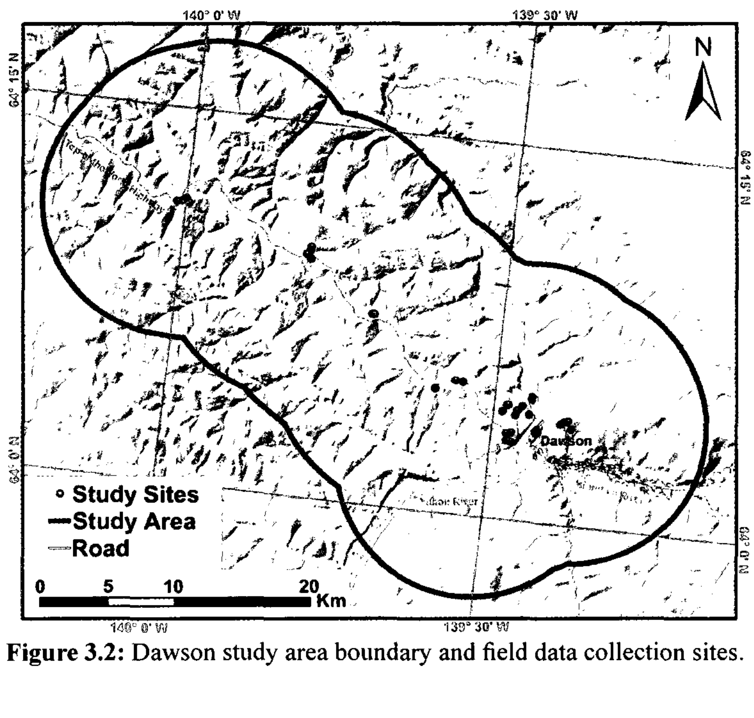


Figure 3.2: Dawson study area boundary and field data collection sites.

3.2 Faro

The Faro study area, Figure 3.3, (at 62° 13.1' N, 133° 8.6' W) is located around the town of Faro and along the Robert Campbell Highway north-west of Ross River. The entire study area encompasses approximately 1410 km^2 and ranges in elevation from about 600 to 2045 m asl. It is located on the boundary between widespread and sporadic permafrost (10 to 90% of the terrain underlain by permafrost) (Heginbottom *et al.*, 1995). Like the Dawson study area, Faro is located within the Central Yukon Basin, but in the south-eastern portion of the region (Figure 3.1) (Wahl *et al.*, 1987). The MAAT (from 1977 to 2000) recorded at the Faro Airport weather station (62° 12' N, 133° 22' W, elevation 716.6 m asl) is -2.2°C, and the mean annual precipitation is 316 mm (Environment Canada, 2009b). This study area is located within the Yukon Plateau-North ecoregion and is characterized by open white spruce, willow, birch, and some lodgepole pine, with black spruce, scrub birch and willow shrubs, and thick moss in areas with poor drainage (Environment Canada, 2009a). A notable characteristic of the vegetation found on slopes during field work in this study area was the presence of very thick organic mats in open white and black spruce forests.

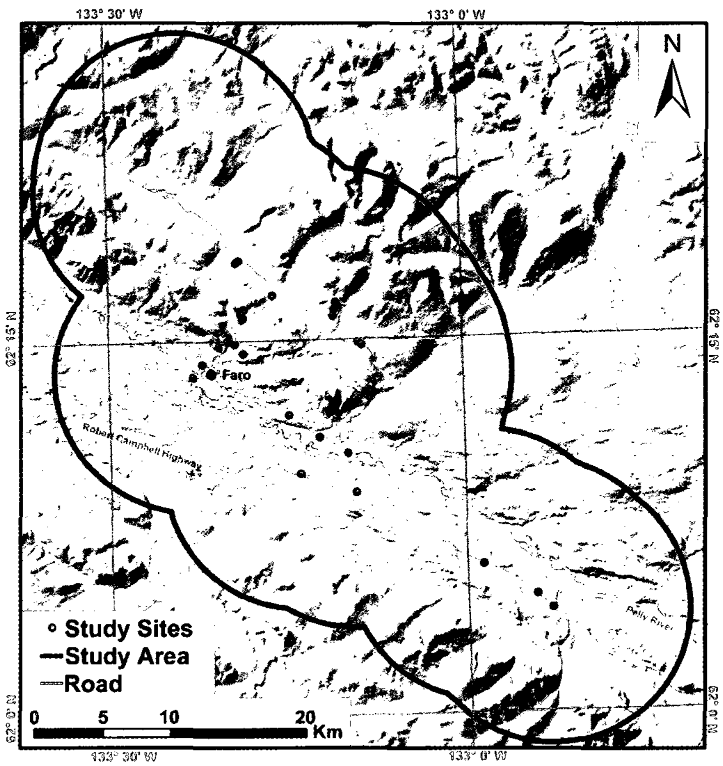


Figure 3.3: Faro study area boundary and field data collection sites.

3.3 KENO

The Keno study area (at 63° 55.3' N, 135° 20.7' W) is located around the town of Keno (Figure 3.4). The entire study area is approximately 820 km^2 and ranges in elevation from about 610 to 2055 m asl. It is within the widespread discontinuous permafrost zone (50 to 90% of the terrain underlain by permafrost) (Heginbottom *et al.*, 1995). It is located in the Ogilvie-Mackenzie Mountains climatological region, characterized by cool summers, moderate winters, and moderate to heavy precipitation (400 - 700 mm per year) (Wahl *et al.*, 1987). The MAAT for three years between 1974 and 1982 recorded at the Keno Hill weather station (63° 56' N, 135° 12' W, elevation 1472.8 m asl) was -5.1°C, and for the four years with complete precipitation data, the mean annual precipitation was 572 mm (Environment Canada, 2009b). More complete and recent data are available for the Mayo Airport weather station (63° 37' N, 135° 52' W, elevation 503.8 m asl), approximately 50 km south-west of the Keno study area. The Mayo Airport weather station recorded a MAAT of -3.1°C and mean annual precipitation of 313 mm between 1971 and 2000 (Environment Canada, 2009b). Keno is in the Yukon Plateau-North ecoregion, like Faro (Environment Canada, 2009a). Based on the data collected in the field, forests in the area are primarily black and white spruce with thick organic mats, while scrub birch and tundra vegetation dominate at higher elevations.

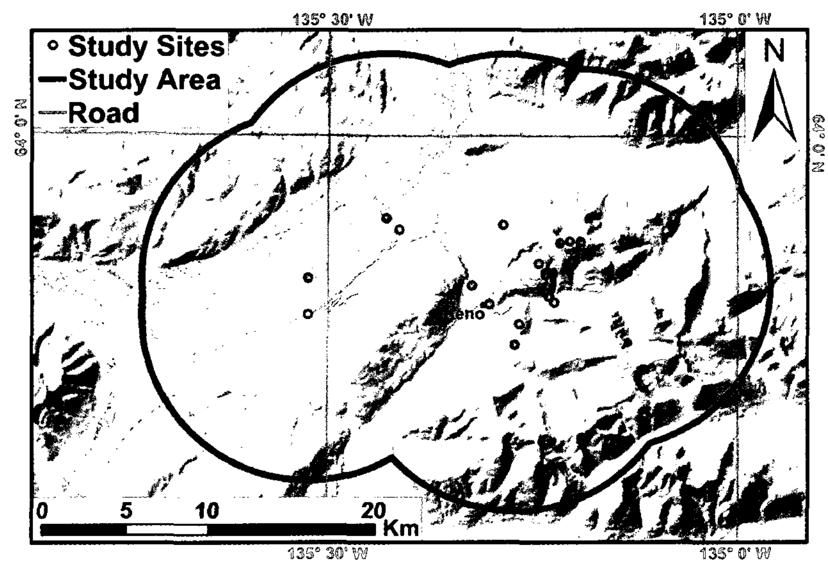


Figure 3.4: Keno study area boundary and field data collection sites.

3.4 Johnson's Crossing

The Johnson's Crossing (Figure 3.5) (60° 34.5' N, 133° 9.5' W) study area is located along the South Canol road approximately 100 km east of Whitehorse and 50 km north-west of the town of Teslin (Figure 3.1). The entire study area is approximately 955 km^2 and ranges in elevation from approximately 670 to 1860 m asl. It is within the sporadic discontinuous permafrost zone (10 to 50% of the terrain underlain by permafrost). It is located within the Pelly-Cassiar Mountains climatological region, characterized by cool summers, moderate winters, and heavy precipitation (500-700 mm per year) (Wahl *et al.*, 1987). The MAAT (from 1971 to 1995) recorded at the now-inactive Johnson's Crossing weather station (60° 29' N, 133° 18' W, elevation 690.4 m asl) was -1.5°C, and the mean annual precipitation was 376 mm (Environment Canada, 2009b). The Teslin Airport weather station, the closest active station (60° 10' N, 132° 44' W, elevation 705 m asl), indicates a MAAT of -1.2°C and mean annual precipitation of 343 mm (Environment Canada, 2009b). Johnson's Crossing is located within the Yukon Southern Lakes ecoregion, which tends to contain open white spruce, lodgepole pine, and trembling aspen forests, with subalpine fir (*Abies lasiocarpa*), dwarf shrubs, and tundra vegetation above the treeline (Environment Canada, 2009a).

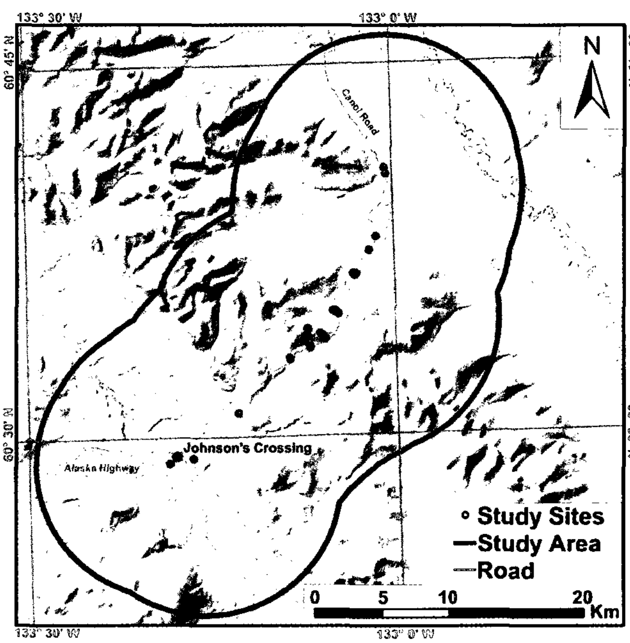


Figure 3.5: Johnson's Crossing study area boundary and field data collection sites.

3.5 Sa Dena Hes Mine

The Sa Dena Hes study area (Figure 3.6) (at 60° 25.5' N, 128° 58' W) is located approximately 45 km north of Watson Lake and along the road to the Sa Dena Hes mine, 0 to 25 km north-east of the Robert Campbell Highway (Figure 3.1). The entire study area is approximately 1050 km^2 and ranges in elevation from approximately 630 to 1565 m asl. It is within the sporadic discontinuous permafrost zone (10 to 50% of the terrain underlain by permafrost) (Heginbottom *et al.*, 1995). Warm summers, cold periods in the winter, and moderate precipitation (400 - 600 mm per year) characterize the Liard Basin climatological region in which Sa Dena Hes is located (Wahl *et al.*, 1987). The MAAT (from 1971 to 2000) recorded at the Watson Lake Airport weather station (60° 06' N, 128° 49' W, elevation 687.4 m asl) is -2.9°C, and the mean annual precipitation is 404 mm (Environment Canada, 2009b). This study area is within the Liard Basin ecoregion, which is characterized by lodgepole pine, white spruce, black spruce, and trembling aspen (Environment Canada, 2009a). At high elevations, above the treeline, low scrub birch and willow shrubs and tundra vegetation dominate.

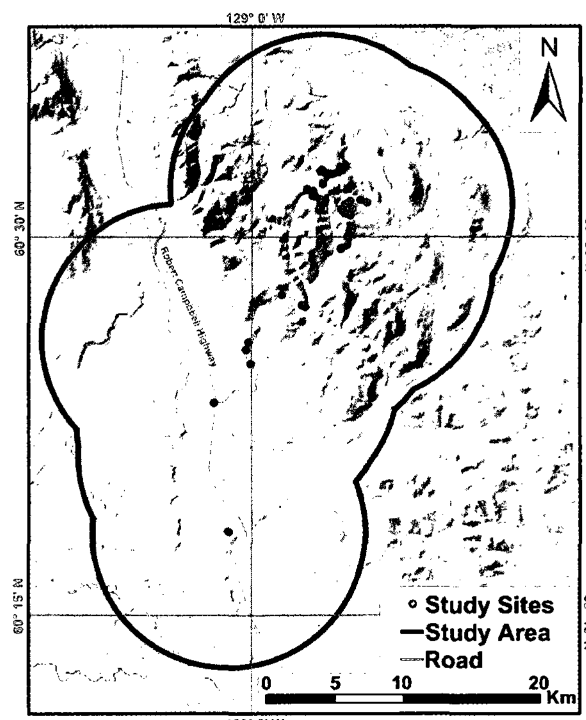


Figure 3.6: Sa Dena Hes study area boundary and field data collection sites.

3.6 Forest Fires

Forest fires that have been recorded within the study areas in the past 70 years are shown in Figure 3.7 (Government of Yukon, Wildland Fire Management, 2008). These fires have certainly disturbed the vegetation cover and probably the permafrost. The balance between climate, vegetation, and permafrost prior to the fire determines the effect on the permafrost (Shur and Jorgenson, 2007). If the climatic conditions are conducive to permafrost development, then the permafrost may recover (Shur and Jorgenson, 2007). However, if vegetation protects the permafrost in unfavourable climatic conditions, then it is possible that the permafrost subsequently thaws and disappears (Shur and Jorgenson, 2007). Regardless of the climate-permafrost conditions prior to the fire, immediately afterward, permafrost degrades (Shur and Jorgenson, 2007; Van Cleve and Viereck, 1983; Viereck, 1972). However, once the organic mat and vegetation cover are re-established, the permafrost can recover as the active layer thins (Van Cleve and Viereck 1983; Viereck, 1972). The switch from a growing active layer to a thinning active layer has been found to occur at some point between 25 to 50 years after the fire (Van Cleve and Viereck, 1983; Viereck, 1972). Therefore, burned areas from fires that have occurred within the past 40 years (since 1970) have been masked out of each study area to account for the time needed for vegetation and permafrost recovery.

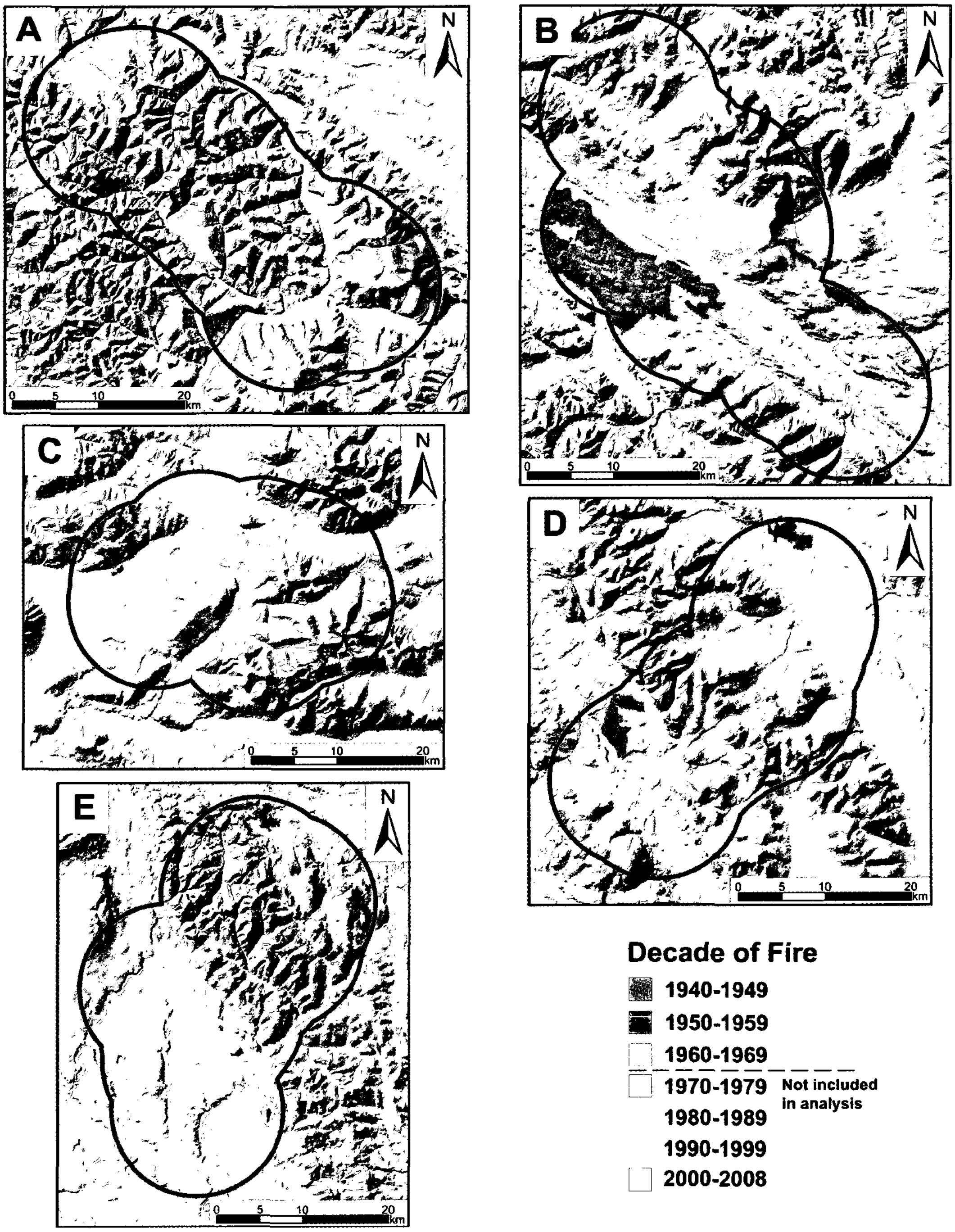


Figure 3.7: Forest fire boundaries and decade of occurrence for each study area. Areas where the forest was burned after 1970 were not included in the analyses.
A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.

4 METHODS

A summary of the procedures used to meet the research objectives of this project is given in Figure 4.1. A total of three DEM-derived variables, four vegetation variables assessed during fieldwork in August 2008, and three vegetation variables extracted from satellite imagery were used in CART models to model permafrost presence or absence. The models were then compared to determine whether vegetation can be a useful factor in the prediction of mountain permafrost distribution.

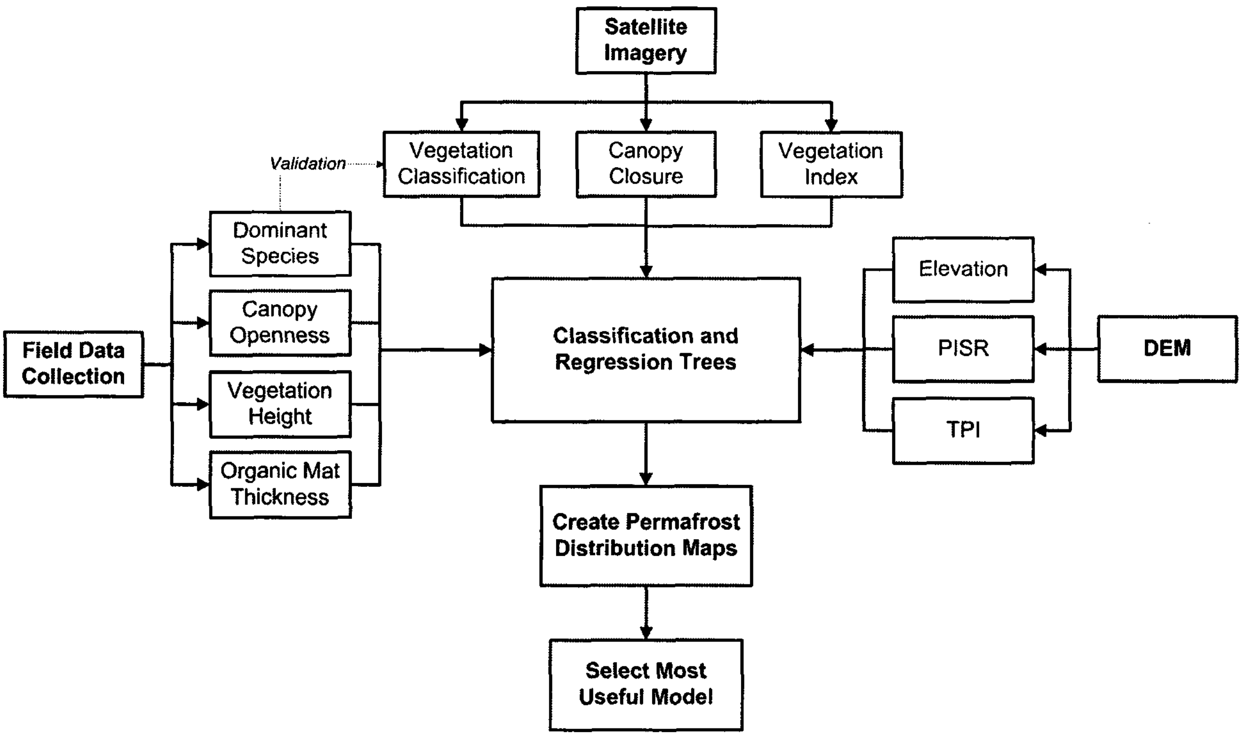


Figure 4.1: Outline of procedures used to complete this research.

4.1 Fieldwork

Fieldwork was undertaken in August 2008. A total of 136 sites were examined: 30 at Dawson, 27 at Faro, 22 at Keno, 27 in Johnson's Crossing, and 30 at Sa Dena Hes. At each site, measurements and observations of vegetation characteristics were recorded at nine points along two intersecting 20 m transects. Although the imagery produces 30 m wide pixels, the 20 m transects were chosen to reduce the effects of error associated with the recorded coordinates. As well, these 20 m transects were more manageable for data collection in the field due to limited time and mobility constraints caused by steep slopes and dense vegetation. The transects followed north-south and east-west bearings to keep the layout the same for all sites (Figure 4.2).

Vegetation characteristics recorded at each of the points included canopy openness, vegetation height, organic mat thickness, and dominant plant species. Canopy openness, vegetation height, and organic mat thickness were all ratio level variables, and the dominant plant species was nominal level. Because decision trees were used for the permafrost distribution model, the levels of measurement did not limit the ability to use any variable.

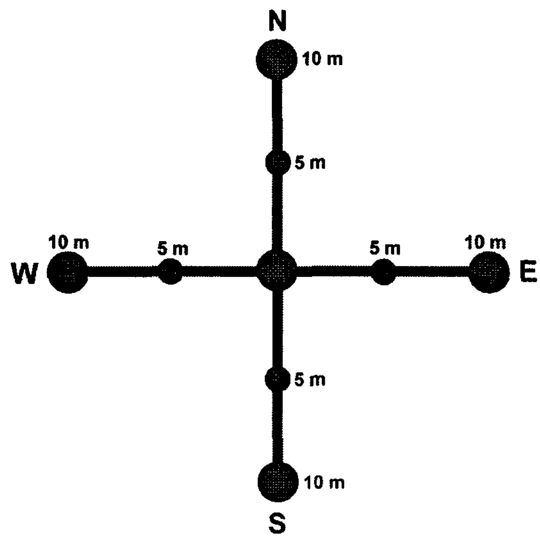


Figure 4.2: Study site layout.

4.1.1 Canopy Openness

Canopy openness was examined at five points at each site: the intersection point and the four end points of the two transects. Several methods for assessing canopy cover were tested by Brown *et al.* (2000), who found the canopy scope method to be the most useful. Consequently, a canopy scope was used to provide a standardized canopy openness assessment for each point. The term "canopy openness" is used here to clarify the meaning of the percentages created: 100% canopy openness means that there is no canopy closure, whereas 0% canopy openness means that there is 100% canopy closure. The canopy scope was selected for this project because it is low-cost, light-weight, and simple to use. The scope was built following the description by Brown *et al.* (2000): 25 holes arranged 3 cm apart in a 5 x 5 grid were drilled into a Plexiglas sheet. A 20 cm string attached to one corner was used to control the distance the tool was held from the user's face. At each measurement point, the canopy scope was centered on the largest gap in the canopy, and the number of dots positioned over the gap was counted. For gaps covering a larger area than the 25 dots of the canopy scope, dots were counted for one section of the gap, then the scope was moved to cover an adjacent but not overlapping section of the gap. This process was repeated until the scope had covered each area of the gap. To move the scope from one section of the gap to another, a visual reference on the edge of the canopy was noted and then the scope was realigned with the reference points. For gaps that exceeded 100 dots, "> 100" was recorded, as this was deemed to be open forest.

The dot count for each point was converted into a percentage based on a polynomial equation created using known values. No dots represents 0% canopy openness (100% canopy cover), 25 dots is equivalent to 30% openness or 70% cover (as suggested by Brown *et al.*, 2000), and 280 dots is equivalent to 100% openness (0% cover). The value of 280 dots was determined by calculating the surface area of a hemisphere with a radius of 20 cm (the distance the scope was held from the user's eyes). This area (2513 cm^2) was then divided by the area covered by the canopy scope

($225\ cm^2$) to determine how many times the canopy scope would fit into that hemisphere. Finally, the answer was multiplied by the number of dots on the canopy scope (25) to find the total number of dots in a fully open hemisphere (approximately 280). The following second-order polynomial equation was found to be the best fit to these three points when the percentage of openness was graphed on the x-axis and the number of dots was placed on the y-axis (Figure 4.3).

$$Y = 0.0281x^2 - 0.0095x \qquad (4.1)$$

In this equation, Y is the number of dots counted and x is the canopy openness percent. Each one hundredth of a percent from 0 to 100 was input as the x value to calculate the number of dots. Only the percentage values that represented whole numbers of dots were extracted and rounded to a tenth of a percent. The table of corresponding percentage and dot values, in Appendix A, was then used to convert the number of dots counted in the field at each point to the percentage of openness. The mean of the five canopy openness percentages was then calculated as the canopy openness for a given field site. The mean was selected for this measurement because the variability in canopy openness values at a given site was low. Homogeneity within a given site with regard to open or closed canopies was a factor when selecting sites, thus the canopy openness values were not affected by outliers.

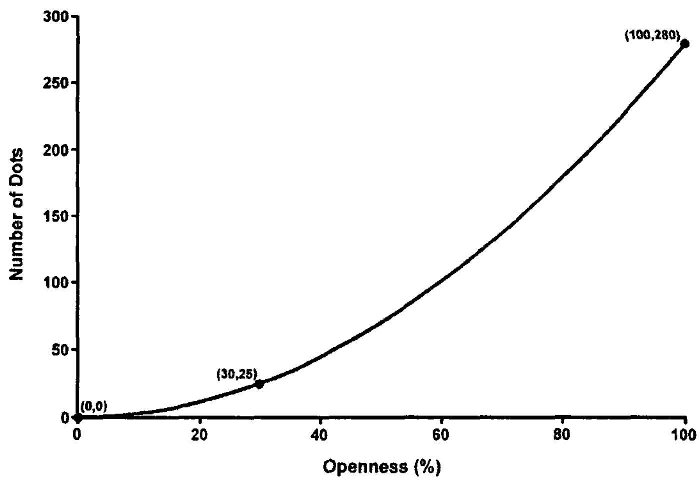


Figure 4.3: Second-order polynomial trend line through known percentage-dot equivalences.

4.1.2 Vegetation Height

Vegetation height was recorded at each of the nine points at a study site. The tallest plant within approximately 2 m of the point was measured and recorded. Shrubs shorter than 2 m were measured with a measuring tape. For taller shrubs and trees, the measurements were made using triangulation (Figure 4.4). The distance from the tree was measured and ranged between 2 m and 7 m due to variations in the density of vegetation. The angle was measured using a clinometer. To calculate the height of the tree or shrub, equation 4.2 was used.

$$H = (D \times \tan\theta) + P \qquad (4.2)$$

In this equation, H is the height of the tree, D is the distance from the tree to the individual taking the measurements, θ is the angle measured in degrees when the clinometer was pointed at the top of the tree, and P is the height at eye level of the individual doing the measurements. The median value of the nine measurements was selected to represent the vegetation height at a given study site. Median was used (as opposed to mean) to account for anomalous values (e.g. from scattered trees in generally low vegetation or open canopy areas).

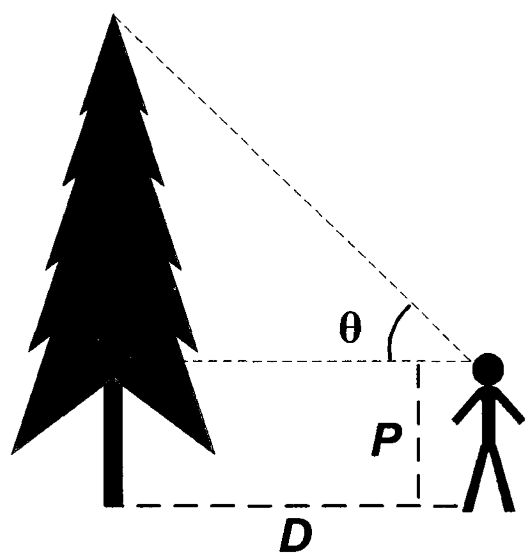


Figure 4.4: Measurements for the tree height equation.

4.1.3 Organic Mat Thickness

Organic mat thickness was measured using a wooden probe at each of the nine points at each study site. The probe (approximately 1 cm in diameter) was gently pushed into the ground until there was slight resistance which corresponded to the top of mineral soil. The measurements were made by the same person in order to maintain consistency for the amount of resistance reached. The mean of the nine recorded values was used to represent the organic mat thickness for each study site. Like the canopy openness values, the organic mat thickness did not vary greatly within a site; thus the data were not affected by outliers, and the mean was representative of site conditions.

4.1.4 Dominant Plant Species

Plant species were recorded as a qualitative assessment of the vegetation type at a particular study site. The different species were identified and recorded at each point along the transects, using Brockman (2001) and MacKinnon *et al.* (1999) as guides. A relative descriptor ("few", "some", "lots") was assigned to describe the frequency of occurrence of each of the species. "Few" indicated that a specific species was located at that site, but the numbers were limited and the species could only be found in isolated locations. "Some" was used to describe species that were located at the majority of points along the transects, but were not the dominant species for the site. "Lots" was used to identify the dominant species at a particular site. This descriptor indicated that multiple plants of a particular species could be found at the majority of the points along or in close proximity to the transects. These subjective descriptors were used to assess the relative dominance of each of the vegetation species at a study site. The classifications were selected by one person to maintain consistency.

4.2 Study Area Delineation

The boundaries for the study areas were created based on the locations of all BTS and ground-truthing points collected in the field from 2006 to 2008. For each study area, the points were imported into ArcGIS 9.2, and a buffer of 10 km around each point was calculated. The dissolve tool was used on the resulting polygons to combine them into a single polygon.

4.3 Satellite Imagery

Satellite imagery was used to examine vegetation characteristics over the study areas. Landsat Thematic Mapper (TM) 5 and Landsat Enhanced Thematic Mapper Plus (ETM+) 7 images (downloaded from US Geological Survey, 2009a) were selected for their spatial and spectral resolution. The specifications of each satellite are displayed Table 4.1. Both platforms produce data at a spatial resolution of 30 m x 30 m, which corresponds to the final permafrost distribution maps. The spectral resolution differs somewhat between the two sensors, but both provide the necessary bands. Both sensors collect data for six bands with 30 m spatial resolution representing the blue, green, and red bands of the visible spectrum and three near- and mid-infrared bands. Thermal bands of coarser spatial resolution and a panchromatic band (from Landsat ETM+7) are also included with the data, but were not used for this research.

Table 4.1: Comparison of specifications for Landsat 5 TM and Landsat 7 ETM+ satellites. (Source: Jensen, 2005)

		Landsat 5 TM	**Landsat 7 ETM+**
Spectral Resolution (μm)	**Band 1**	0.45 – 0.52	0.45 – 0.515
	Band 2	0.52 – 0.60	0.525 – 0.605
	Band 3	0.63 – 0.69	0.63 – 0.69
	Band 4	0.76 – 0.90	0.75 – 0.90
	Band 5	1.55 – 1.75	1.55 – 1.75
	Band 7	2.08 – 2.35	2.08 – 2.35
Spatial Resolution		30 m x 30 m	30 m x 30 m
Altitude		705 km	705 km
Swath Width		185 km	185 km
Inclination		98.2°	98.2°
Revisit		16 days	16 days

4.3.1 Image Selection

The most recent, cloud-free (over the study area), growing season image from either platform was selected for each study area. Images acquired early in August were preferred as best representing the vegetation. Due to the failure of the Scan Line Corrector (which accounts for and corrects gaps created due to the forward movement of the satellite) with Landsat 7 in May 2003, more recent images contain gaps that require multiple images to correct (US Geological Survey, 2009b). The correction process proved to be too lengthy for this project, and the results were inadequate for consistent comparisons between bands. Thus it was decided that images prior to the 2003 malfunction would be acceptable if no Landsat 5 image was available for a study area. Changes in vegetation cover that would affect the results of this research include forest fires or large-scale clear-cutting. Fires that occurred within the study areas since the oldest images were collected were masked out of the analysis; therefore the differing years selected should not significantly affect the vegetation data collected for this research. The date and platform for each image for each study area are listed in Table 4.2.

Few cloud-free images exist for Dawson and Keno in the summer months. The late August image selected for Keno includes some senescence of the deciduous vegetation, which could pose challenges for the vegetation classification. The early July image of Dawson was selected to better represent the vegetation than the late August or early September options. This image is not cloud-free, but the clouds are not prominent in the study area itself and have been masked out for the project. The two images for the Faro area were corrected separately and subsequently mosaiked together to create a nearly seamless image of the study area.

Table 4.2: Images used for each study area.

Study Area	Sensor	Path	Row	Acquisition Date
Dawson	Landsat 5	064	015	July 3, 2008
Faro*	Landsat 7	058	017	August 15, 2001
	Landsat 7	059	016	August 1, 1999
Keno	Landsat 5	061	015	August 31, 2008
Johnson's Crossing	Landsat 7	057	018	August 3, 1999
Sa Dena Hes	Landsat 5	055	018	August 5, 2008

**Two images selected because the study area is located at the edge of both images*

4.3.2 Pre-processing

Geometric correction of all images was completed prior to distribution by the USGS. For each image, all bands were imported into PCI Geomatica 10.1 and were then merged into a single file for each study site using the Data Merge tool. The UTM zone, extents, and pixel size were set according to the information provided in the metadata of the images, and the bilinear resampling method was used to reproject the images. The images were then clipped to a rectangle approximately 10 km larger (on all sides) than the defined study area boundaries to avoid errors caused by edge effects. This distance produced a much larger area than needed and, after all calculations were complete, the data were clipped to the boundaries of the study areas. Prior to analysis, radiometric correction was performed to reduce the effects of haze and shadows in the imagery.

4.3.3 Radiometric Correction

The initial step for the radiometric correction was a simple histogram adjustment. This adjustment moves each histogram so that the minimum value is zero (Jensen, 2005). Because the offset for each band is different, this adjustment balances the histograms relative to each other. This adjustment effectively reduces haze in the images by balancing the high offset in the blue band (due to atmospheric scattering) with the low offset in the mid-infrared bands. This correction does not, however, correct for topographical effects on brightness values (Jensen, 2005).

Shadows created by changes in topography can be corrected using an algorithm that accounts for PISR based on slope angle and slope aspect:

$$L_H = L_T - (\cos(i)m \times 100) - b + \overline{L_T} \quad (4.3)$$

where L_H is the corrected brightness value, L_T is the uncorrected brightness value, *m* and *b* are the slope and y-intercept of the linear regression equation between the uncorrected image and ($\cos(i)m \times 100$), and $\overline{L_T}$ is the mean of the brightness values for the uncorrected image (Meyer *et al.*, 1993).

Several other similar corrections (e.g., a simple cosine correction) also exist. The method used was selected for its simplicity and because it has been found in the literature to be the most useful (Meyer *et al.*, 1993; Teillet *et al.*, 1982).

A DEM for each study area was clipped from a DEM of the entire Yukon Territory (see section 4.4). The DEMs were then reprojected into the same coordinate system as the imagery for each study area (UTM, datum: WGS 1984). Each DEM was then imported into Geomatica 10.1, clipped to the same extent as the study area image, and merged into the study area image file. To calculate ($\cos(i)m \times 100$), first the angle calculation tool within Geomatica 10.1 (ANG) was used with the following parameters: the upper-left corner location was copied from the metadata files with the images, and the distance to the light source was calculated using a calculator available online, the parameters for which were also found in the metadata files with the exception of acquisition time, which was set to 12:00:00 UTC when not specified (Walker, 1995). The angles calculated were then converted from degrees into radians using the raster calculator tool. Finally, using the raster calculator tool, ($\cos(i)m \times 100$), was calculated.

The regression tool was used, comparing the ($\cos(i)m \times 100$) values to those of the histogram adjusted image, to find m and b for equation 4.3 (Jensen, 2005; Meyer *et al.*, 1993; Teillet *et al.*, 1982). The mean value was found by examining the statistics included with the histograms associated with each band. The final equation was then calculated using Easi Modeller.

Faro required two images to cover the entire study area. In this case, a mask was created for each image to represent the area it would cover in the final mosaic. Radiometric corrections were performed separately for each image under the masks in the same manner as for the other study areas. The two images were then merged together using Easi Modeller where only image values were placed into the output image where the mask value for that image was 1.

4.3.4 NDVI

A NDVI was calculated for each study area using the following equation:

$$NDVI = \frac{(NIR_{(Band\ 4)} - Red_{(Band\ 3)})}{(NIR_{(Band\ 4)} + Red_{(Band\ 3)})} \quad (4.4)$$

where NIR and Red represent the near infrared band 4 and red band 3 respectively in Landsat imagery (Rouse *et al.*, 1974). This equation examines the relationship between the near infrared band and the red band, which can be used as an indicator of vegetation greenness (Rouse *et al.*, 1974). The spectral signature of green vegetation shows very low reflectance in the red wavelengths, but very high reflectance in the near infrared wavelengths (Rouse *et al.*, 1974; Jensen, 2005). Therefore a ratio between these two bands shows the intensity of this difference. NDVI is a modification of the simple ratio that produces a range of values between -1.0 and 1.0. Lower values (zero and below) represent water, roads, and other unvegetated areas. Values closer to 1.0 represent healthier vegetation. This vegetation index is one of the simplest indices to calculate and has been used in the past to represent vegetation for the purpose of modelling permafrost distribution (e.g. Leverington and Duguay, 1996; Ødegård *et al.*, 1999).

4.3.5 Vegetation Classification

Vegetation classifications were created using the Landsat images for each study area using a combination of unsupervised and supervised classifications undertaken with PCI Geomatica 10.1. First, two unsupervised classifications (one using the K-Means method, the other using ISOData) were created for each study area. For both K-Means and ISOData most parameters were left as the defaults, with the exception of the maximum number of classes, which was changed to 50; for ISOData, the minimum number of classes was changed to 25 and the desired number of classes was changed to 50. K-Means and ISOData are similar methods for classifying data. The primary difference between the methods is that the number of classes does not change between iterations with the K-Means method, but does fluctuate between iterations with ISOData (Lillesand *et al.*, 2008). Both methods were tested to determine which would be best for classifying vegetation in each study area and in all five areas as a whole.

Each field site was assigned to a vegetation class that represented the dominant vegetation type at that site (based on observations made in the field). Each site was labelled as one of four classes: tundra, shrub, broadleaf trees, or needle-leaf trees. This list of four was reduced from 16 to improve the accuracy of the classification because, in preliminary tests, the software could not adequately differentiate between shrub heights, needle-leaf species, or open vs. closed canopy of individual species in the imagery. The field sites, a point layer in the GIS, were then overlaid on each of the two classifications. The automatically created class that coincided with each field site was labelled according to the vegetation class of that site. Water and bare ground classes were defined using clearly visible features such as lakes, roads, and mines.

The newly labelled classes were then aggregated into seven classes: water, bare ground, tundra, shrub, broadleaf trees, needle-leaf trees, and undefined areas. The undefined class represented all classes created by the unsupervised classification that were not represented by field sites and could not be easily defined as water or bare ground.

Subsequent to the definition of the classes created through the unsupervised classification, a supervised classification was created to fill in the undefined areas. Training sites for the supervised classification were drawn based on the defined areas from the aggregation of the unsupervised classification. The maximum likelihood classifier was used to perform the supervised classification.

An accuracy assessment of the final classification was completed using the 2008 field data as well as some points from the summer 2007 field work. The more accurate of the K-Means or ISOData methods was selected for each area based on the accuracy assessment.

4.3.6 Canopy Closure Classifications

Classifications of canopy closure were created from the imagery using a similar method to that used for the vegetation classification. For these classifications, only two classes (open and closed canopy) were defined using the field sites, in addition to a water class. Using these classes, the same procedure as the one used for the vegetation classes was followed.

4.4 DEM

Three variables that either are implicit to a DEM or can be derived from one have already been found to be useful for predicting the distribution of mountain permafrost: elevation, PISR, and topographic position index (TPI) values. These variables were also included in the models for this research. Elevation values were collected directly from a DEM of the Yukon created by the Yukon Geological Survey (2006) with a resolution of 30 m x 30 m. For calculating PISR and TPI, the DEM was clipped to rectangular boxes approximately 10 km larger (on each side) than the extent of each study area boundary.

4.4.1 PISR

Previous research involving mountain permafrost modelling has found PISR (or other proxy variables like equivalent latitude or slope aspect) to be particularly useful (e.g., Gruber and Hoelzle, 2001; Leverington and Duguay, 1996; Morrissey *et al.*, 1986; Lewkowicz and Bonnaventure, 2008). For this research, PISR was calculated using the Solar Radiation tool available within the Spatial Analyst extension in ArcMap 9.2. PISR values were then extracted for each field site within each study area.

The parameters selected for these models are listed in Table 4.3. Cloud cover percentages were estimated by subtracting the ground surface temperature from the air

Table 4.3: Parameter values for the solar radiation models.

Parameter	Value
Centre Latitude	Calculated automatically
Sky Size	400 x 400 cells
Time Configuration	Multiple Days in a Year
Start Date	May 15
End Date	September 30
Day/Hour Interval	Default
Topographic Parameters	Default
Zenith Divisions	16
Azimuth Divisions	16
Diffuse Model Type	Standard Overcast Sky
Diffuse Proportion	Cloud Cover Percent
Transmittivity	Default

temperature for one-hour intervals throughout the snow-free period in 2007 (May 15 to September 30) at two logger sites within each study area. These differences were then graphed for every 24 hours within the specified period and, based on the shape of the graph, the given day was assigned a cloud cover percent of 0%, 50%, or 100%. Days with large differences between the air and ground temperatures indicated that the ground surface was being affected by direct solar radiation. Days with little difference between air and ground temperatures indicated that there was little direct solar radiation affecting the temperatures and were assumed to be mainly cloudy. Days for which the graphs vary from hour to hour were assigned 50% cloud cover. The mean of the percentages for the two logger locations was set as the cloud cover percent for the study area.

4.4.2 TPI

A Topographic Position Index (TPI) was used to account for areas potentially affected by cold air drainage, where permafrost appears or is common in valleys while none appears or is less common at higher elevations. The index was calculated using an extension by Jenness Enterprises for ArcView 3.x (Jenness, 2006). Using the DEM clipped for each study area, the TPI tool was run using a moving circular window with a radius of 1000 m. TPI values represent how the elevation of each cell relates to the mean of the surrounding cells (defined by the neighbourhood parameters). Negative values indicate lower points (e.g., valleys), positive values indicate higher points (e.g., mountain tops or ridges), and values of zero indicate flat terrain or long slopes (Jenness, 2006; Weiss, 2001).

4.5 Additional Variables

In addition to the DEM-derived variables, two other non-vegetation variables were included: latitude and longitude. These variables were calculated using the UTM coordinates of each point. These coordinates were converted into decimal degrees using UTM Conversion Tool (Dutch, 2005).

4.6 Analysis

4.6.1 Variable Comparisons

Boxplots were used to examine within-variable differences with regard to permafrost presence or absence. These plots were created with presence or absence of permafrost on the x-axis and the variable on the y-axis.

All variables were tested for correlation using the Pearson correlation coefficient calculated with PASW Statistics 17.0 (SPSS Inc., 2009b). Significant correlations between variables (at the 0.05 level) were used to explain the results of the decision tree analysis and to examine how the variables relate to each other.

4.6.2 Decision Tree Classification

Because this research involved both quantitative and qualitative data, logistic regression could not be used while incorporating all variables. Instead, regression tree models were created using PASW Statistics 17.0 to model permafrost distribution using both nominal and interval or ratio level vegetation and non-vegetation variables (SPSS Inc., 2009a). These models are also useful because they can take into account some of the more complicated relationships between the variables and permafrost presence or absence. In some cases, a variable may only be useful under certain conditions or the relationship between the variable and permafrost may not be linear. Decision trees can help account for these types of relationships (StatSoft, 2007). Each set of models contains one model based on all study sites in the five study areas and separate models for the sites from each of the study areas.

The models were created using 102 points, and they were tested using the remaining randomly selected 34 (25%) points. The same set of test sites was used for each tree. Two additional sets of trees were created with two different test samples to assess whether the original test sample was representative of the dataset. A random number generator was used to select sets of 34 numbers between 1 and 136 (the total number of sites). Then the sites with identification numbers corresponding to the 34 numbers (the testing set) were coded with the number 2 while all others (the training set) were coded with the number 1 in a field on the spreadsheet. This field was then used as to differentiate between training and testing sites.

The models were created automatically using the same parameters to maintain consistency. To prevent over-fitting (creating so many splits that the tree no longer represents the population and is, instead, extracting single outlier sites or small groups of outliers), pruning was undertaken based on a set minimum number of samples required for a child node to be produced, and on a set minimum improvement for a split to occur (Breiman *et al.*, 1998; Ripley, 1996; StatSoft, 2007). The minimum number of samples required in a child node was set at three to prevent single sites from being extracted. The minimum improvement value was calculated by the software after each split in the tree using a Gini index to measure impurity of the end nodes (Breiman *et al.*, 1998; Ripley, 1996; SPSS Inc., 2009a; StatSoft, 2007). The change in impurity or improvement is the difference between the impurity of the node prior to the split and the weighted average of the impurities of the nodes after the split. Equation 4.5 was used to calculate improvement.

$$\textit{Impurity} = \sum(T_i \times \mathrm{Gini}_i) \;/\; n \qquad (4.5)$$

where T_i is the number of sites at node *i*, *n* is the total number of sites used for the tree, *i* is each end node in the tree, and Gini_i is calculated using equation 4.6.

$$\mathrm{Gini}_i = 1 - \sum(i_j / T_i)^2 \qquad (4.6)$$

where i_j represents the number of sites of category *j* at node *i* (Breiman *et al.*, 1998; Ripley, 1996; SPSS Inc., 2009a; StatSoft, 2007).

The minimum improvement value was set at 0.009 for the trees based on all sites. It was selected to reduce over-fitting of the data and was based on examination of unpruned trees. For the models for individual study areas, the default of 0.0001 for the minimum change in improvement was used as the criterion for the minimum number of sites in a child node. This level sufficiently pruned the trees.

An outline of the models created and the variables included in each one is given in Table 4.4. A DEM-Based model containing no vegetation data (only elevation, PISR, and TPI) was used as a base case to allow subsequent improvements associated with the addition of vegetation variables to be assessed. Two vegetation-only models were created for the same purpose as the DEM-Based model. The Field-Based model contained only vegetation data collected in the field (dominant plant species, organic mat thickness, vegetation height, and canopy openness). The Imagery-Based tree included all vegetation variables extracted from the satellite imagery (NDVI, vegetation classification, and canopy closure).

Table 4.4: Decision tree models and input variables.

Name	Description	Input Variables
DEM-Based	DEM-derived data only	Elevation, PISR, TPI
Field-Based	Field data only	Dominant Species, Organic Mat Thickness, Vegetation Height, Canopy Openness
Field and DEM-Based	Field data and DEM-derived data	Elevation, PISR, TPI, Dominant Species, Organic Mat Thickness, Vegetation Height, Canopy Openness
Imagery-Based	Imagery-derived data only	NDVI, Vegetation Classification, Canopy Closure Classification
Imagery and DEM-Based	Imagery-derived data and DEM-derived data	Elevation, PISR, TPI, NDVI, Vegetation Classification, Canopy Closure Classification

Two additional models were created using the variables from both the base model and each of the vegetation-only models. The Field and DEM-Based tree examined how much the field data could improve the classification of permafrost presence or absence and included elevation, PISR, TPI, dominant plant species, organic mat thickness, vegetation height, and canopy openness. The Imagery and DEM-Based tree tested how much the vegetation data extracted from the satellite imagery could improve the DEM-

Based model and included elevation, PISR, TPI, NDVI, vegetation class, and canopy closure.

Models were compared based on the most important variables, the complexity of the models, the overall accuracy for both the training and testing datasets, the separate accuracies for the permafrost probable and permafrost improbable classes, and the overall improvement (difference between the impurity prior to any splits and the impurity of the final model).

4.6.3 Maps

Maps of predicted permafrost distribution for each study area were created in PCI Geomatica 10.1 using conditional statements written in EASI Modeller. Only the tree models that included variables for which complete coverage of the study areas was available (the DEM-Based tree and the two imagery-based trees) were used for these maps. For each study area, a map was created using the tree models with all the data and the models created using only the data from the individual study area.

The actual study area boundaries were used to clip each permafrost distribution map, and the polygons representing fires recorded since 1970 were masked out to remove potentially incorrect predictions in areas that have not sufficiently recovered from fire. Finally, the predicted extent of permafrost within each study area was calculated by multiplying the number of pixels classified as permafrost by the size of each pixel (900 m^2). The percentage of permafrost as predicted by the CART models within each study area (area covered by permafrost divided by the total area) was also calculated and compared to the more generalized permafrost map created by Heginbottom *et al.* (1995). The map produced using the model with the highest accuracy was labelled as the best map.

5 RESULTS

The table of all data for the ten variables used in the CART models is included in Appendix B.

5.1 Classifications of Satellite Imagery

5.1.1 Vegetation Classification

Both the K-Means method for the initial unsupervised classification and the ISOData method were used for the classification of the satellite imagery into vegetation classes. The overall, producer's, and user's accuracies for each study area and for each method are displayed in Tables 5.1 and 5.2.

The Dawson image was classified better using the K-Means method (79.1% overall accuracy) than using the ISOData method (73.3% overall accuracy) (Tables 5.1 A and 5.2 A). The Faro image was classified equally well using both methods (71.3% overall accuracy) (Tables 5.1 B and 5.2 B). The ISOData method was better for classifying the Keno image, with an overall accuracy of 66.7% (as opposed to the 61.1% overall accuracy produced by the K-Means-based classification) (Tables 5.1 C and 5.2 C). The Johnson's Crossing image was classified better using the K-Means method (79.2% overall accuracy) (Tables 5.1 D and 5.2 D). Sa Dena Hes was classified better using the ISOData method (71%) (Tables 5.1 E and 5.2 E).

Overall, the two methods proved to have similar accuracies (each method was better than the other in two study areas and they were the same at one study area). However, the K-Means method had higher overall user's accuracy (lower probability of pixels being incorrectly included in a given class) and producer's accuracy (lower probability of pixels being incorrectly omitted from a given class) (Jensen, 2005). Additional analyses confirmed the superiority of the K-Means method for the subsequent regression tree analyses. As well, two initial tests of tree models combining all variables (elevation, PISR, TPI, NDVI) with the K-Means-based classification and the

ISOData-based classification were carried out (Appendix C). Both models had the same overall accuracies for the training data, but the K-Means method had a slightly higher overall accuracy with the testing data.

Table 5.1: K-Means-based vegetation classification accuracy assessment for each study area.
A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.

A

Class	Producer's Accuracy	User's Accuracy
Water	83.3%	100.0%
Bare Ground	100.0%	71.0%
Tundra	87.5%	77.8%
Shrub	85.7%	85.7%
Broadleaf	37.5%	60.0%
Needle Leaf	53.3%	80.0%
Overall Accuracy: 79.1%		

95% Confidence Interval

B

Class	Producer's Accuracy	User's Accuracy
Water	100.0%	100.0%
Bare Ground	95.2%	90.9%
Tundra	100.0%	37.5%
Shrub	72.7%	47.1%
Broadleaf	8.3%	33.3%
Needle Leaf	38.5%	50.0%
Overall Accuracy: 71.3%		

95% Confidence Interval

C

Class	Producer's Accuracy	User's Accuracy
Water	100.0%	100.0%
Bare Ground	90.0%	56.3%
Tundra	80.0%	66.7%
Shrub	28.6%	40.0%
Broadleaf	66.7%	25.0%
Needle Leaf	33.3%	100.0%
Overall Accuracy: 61.1%		

95% Confidence Interval

D

Class	Producer's Accuracy	User's Accuracy
Water	88.9%	100.0%
Bare Ground	100.0%	75.0%
Tundra	100.0%	100.0%
Shrub	66.7%	75.0%
Broadleaf	0.0%	0.0%
Needle Leaf	64.3%	100.0%
Overall Accuracy: 79.2%		

95% Confidence Interval

E

Class	Producer's Accuracy	User's Accuracy
Water	100.0%	100.0%
Bare Ground	91.7%	52.4%
Tundra	41.7%	100.0%
Shrub	80.0%	33.3%
Broadleaf	40.0%	50.0%
Needle Leaf	42.9%	69.2%
Overall Accuracy: 61.3%		

95% Confidence Interval

Table 5.2: ISOData-based vegetation classification accuracy assessment for each study area.
A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.

A

Class	Producer's Accuracy	User's Accuracy
Water	83.3%	100.0%
Bare Ground	100.0%	66.7%
Tundra	87.5%	58.3%
Shrub	61.9%	92.9%
Broadleaf	50.0%	50.0%
Needle Leaf	46.7%	77.8%
Overall Accuracy: 73.3%		

95% Confidence Interval

B

Class	Producer's Accuracy	User's Accuracy
Water	100.0%	100.0%
Bare Ground	95.2%	100.0%
Tundra	100.0%	42.9%
Shrub	54.5%	46.2%
Broadleaf	50.0%	37.5%
Needle Leaf	15.4%	50.0%
Overall Accuracy: 71.3%		

95% Confidence Interval

C

Class	Producer's Accuracy	User's Accuracy
Water	100.0%	100.0%
Bare Ground	80.0%	100.0%
Tundra	60.0%	50.0%
Shrub	57.1%	47.1%
Broadleaf	0.0%	0.0%
Needle Leaf	58.3%	58.3%
Overall Accuracy: 66.7%		

95% Confidence Interval

D

Class	Producer's Accuracy	User's Accuracy
Water	88.9%	100.0%
Bare Ground	92.3%	60.0%
Tundra	100.0%	100.0%
Shrub	37.5%	50.0%
Broadleaf	0.0%	0.0%
Needle Leaf	64.3%	81.8%
Overall Accuracy: 72.9%		

95% Confidence Interval

E

Class	Producer's Accuracy	User's Accuracy
Water	85.7%	100.0%
Bare Ground	83.3%	83.3%
Tundra	75.0%	75.0%
Shrub	40.0%	25.0%
Broadleaf	60.0%	50.0%
Needle Leaf	66.7%	77.8%
Overall Accuracy: 71.0%		

95% Confidence Interval

5.1.2 Canopy Closure Classification

The K-Means and the ISOData methods for the initial unsupervised classification were both tested for the canopy cover classification of the satellite imagery. The overall, producer's, and user's accuracies for each study area are displayed in Tables 5.3 for the K-Means-based classification and 5.4 for the ISOData-based classification.

The Dawson image had a higher overall accuracy with the K-Means method (83.3% overall accuracy) than the ISOData method (80% overall accuracy) (Tables 5.3 and 5.4 A). The ISOData-based method (81.5% overall accuracy) was much better than the K-Means-based method (59.3% overall accuracy) for the Faro image (Tables 5.3 and 5.4 B). For the Keno image, the K-Means method (59.1% overall accuracy) was better than the ISOData-based method (54.5% overall accuracy) (Tables 5.3 and 5.4 C). For the Johnson's Crossing image, the ISOData-based method (74.1% overall accuracy) was better than the K-Means-based method (66.7% overall accuracy) (Tables 5.3 and 5.4 D). For the Sa Dena Hes image, the K-Means-based method (80% overall accuracy) was better than the ISOData-based method (76.7% overall accuracy) (Tables 5.3 and 5.4 E).

The K-Means method had a better overall accuracy than the ISOData-based method in the majority of the study areas. Thus, the K-Means-based classifications were used for all decision tree models that involved a canopy closure classification.

Table 5.3: K-Means-based canopy closure classification accuracy assessment for each study area.
A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.

A

Class	Producer's Accuracy	User's Accuracy
Open	100.0%	80.0%
Closed	50.0%	100.0%
Overall Accuracy: 83.3%		

95% Confidence Interval

B

Class	Producer's Accuracy	User's Accuracy
Open	38.9%	100.0%
Closed	100.0%	45.0%
Overall Accuracy: 59.3%		

95% Confidence Interval

C

Class	Producer's Accuracy	User's Accuracy
Open	63.2%	85.7%
Closed	33.3%	12.5%
Overall Accuracy: 59.1%		

95% Confidence Interval

D

Class	Producer's Accuracy	User's Accuracy
Open	76.9%	62.5%
Closed	57.1%	72.7%
Overall Accuracy: 66.7%		

95% Confidence Interval

E

Class	Producer's Accuracy	User's Accuracy
Open	86.7%	76.5%
Closed	73.3%	84.6%
Overall Accuracy: 80.0%		

95% Confidence Interval

Table 5.4: ISOData-based canopy closure classification accuracy assessment for each study area. *A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.*

A

Class	**Producer's Accuracy**	**User's Accuracy**
Open	100.0%	76.9%
Closed	40.0%	100.0%
Overall Accuracy: 80.0%		

95% Confidence Interval

B

Class	**Producer's Accuracy**	**User's Accuracy**
Open	77.8%	93.3%
Closed	88.9%	66.7%
Overall Accuracy: 81.5%		

95% Confidence Interval

C

Class	**Producer's Accuracy**	**User's Accuracy**
Open	63.2%	80.0%
Closed	0.0%	0.0%
Overall Accuracy: 54.5%		

95% Confidence Interval

D

Class	**Producer's Accuracy**	**User's Accuracy**
Open	76.9%	71.4%
Closed	71.4%	76.9%
Overall Accuracy: 74.1%		

95% Confidence Interval

E

Class	**Producer's Accuracy**	**User's Accuracy**
Open	86.7%	72.2%
Closed	66.7%	83.3%
Overall Accuracy: 76.7%		

95% Confidence Interval

5.2 Permafrost Present/Absent Data Comparison

Box plots were created to show the differences between permafrost present and permafrost absent data for each variable. The results of the independent samples difference of means t-test comparing the permafrost absent and permafrost present datasets for each variable corresponded to those of the box plots. The *p*-values for each variable are shown in Table 5.5. These analyses provide some insight into why the tree models included certain variables.

Table 5.5: *p*-Values from t-tests for permafrost present and absent data for each variable. *Significant differences in bold (α=0.05).*

Variable	*p*-Value
Elevation	0.074
PISR	**0.000**
TPI	0.613
NDVI	**0.029**
Canopy Openness	0.121
Vegetation Height	0.093
Organic Mat Thickness	**0.004**

For the elevation box plot (Figure 5.1 A), the range of permafrost absent sites extended within the permafrost present range, but the median of the permafrost absent sites was, surprisingly, slightly higher than the permafrost present median. The two datasets also had opposing skews (the permafrost absent dataset appears to have a very slight negative skew while the permafrost present dataset appears to have a slight positive skew). The t-test for elevation indicates that the means are not significantly different between the permafrost present and permafrost absent data.

In the PISR box plot (Figure 5.1 B), the range of the permafrost absent data also extended into the range of the permafrost present data, but the concentration of the permafrost absent data was higher. As well, the permafrost absent median was visibly higher than the permafrost present median, and the two datasets had opposing skews like the elevation data (but more defined). The t-test indicates that mean PISR is significantly different from permafrost present and permafrost absent data.

The TPI box plot (Figure 5.1 C) showed that both the permafrost absent and permafrost present datasets were very similar except for a slightly greater variance in the permafrost present data. The t-test for TPI showed that the means for permafrost present and permafrost absent data were not significantly different.

In the NDVI box plot (Figure 5.1 D), the range of the permafrost absent data extends slightly higher than that of the permafrost present data, but the majority of the data do overlap. Based on the t-test, however, NDVI values had a significantly lower mean for permafrost present sites than for permafrost absent sites.

The box plots for both the canopy openness data and the vegetation height (Figures 5.1 E and F, respectively) show that both the permafrost present and the permafrost absent datasets are very similar except that the variance of the permafrost absent data is slightly greater. For both variables, the means of permafrost present and permafrost absent were not significantly different based on the results of the t-tests.

The box plot for the organic mat thickness (Figure 5.1 G) shows that the median of the permafrost absent dataset was visibly lower and there was less variance than the permafrost present data. The t-test also indicated that the means for permafrost present and permafrost absent data were significantly different.

Overall, only PISR and, to a lesser degree, NDVI, and organic mat thickness show strong and significant differences between the permafrost present and permafrost absent sites (Table 5.5).

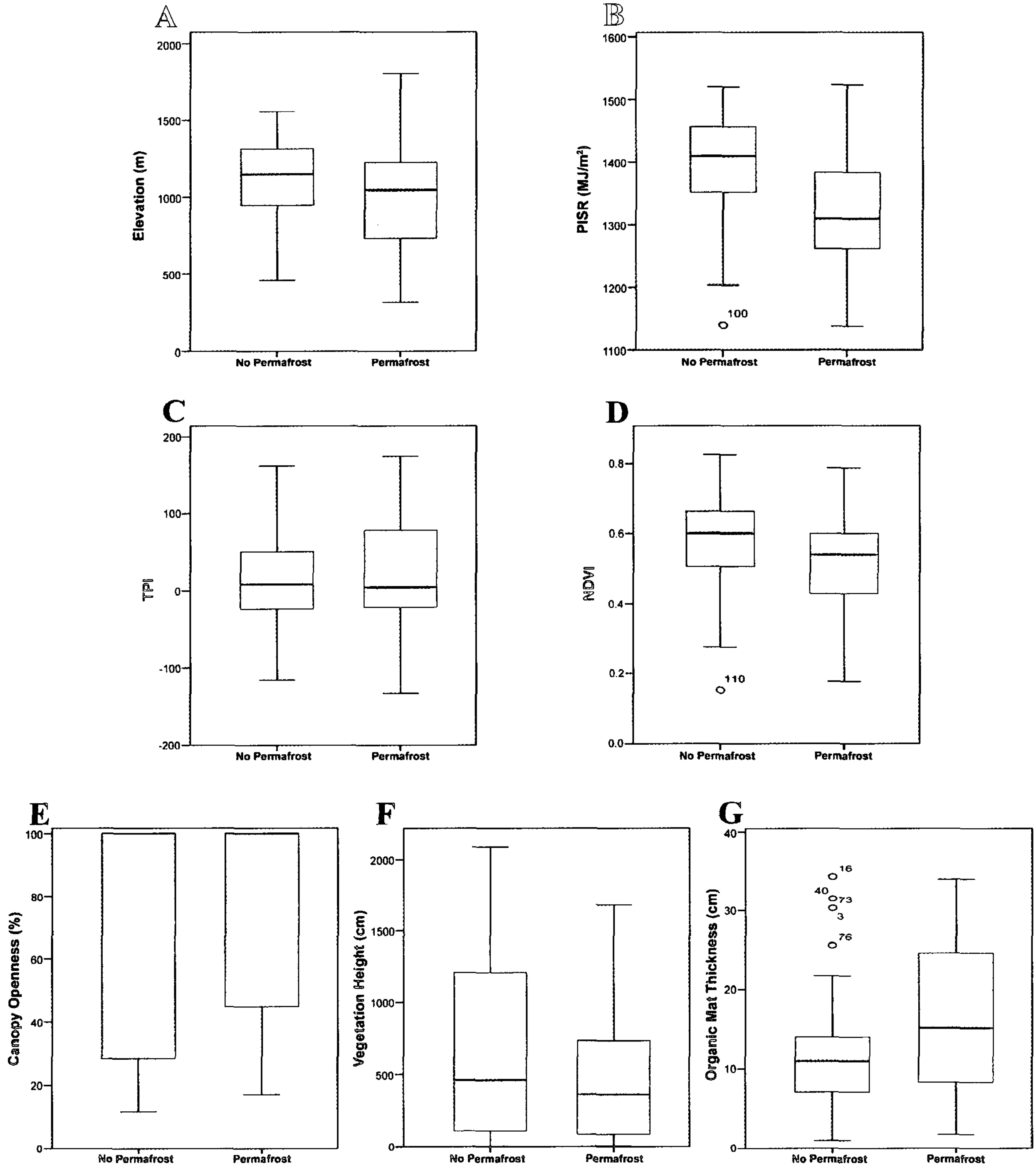


Figure 5.1: Boxplots comparing permafrost present and absent data for each variable.
The middle line in each box represents the median, and the upper and lower extents of each box represent the upper and lower quartiles of the data. A: Elevation; B: PISR; C: TPI; D: NDVI; E: Canopy Openness; F: Vegetation Height; G: Organic Mat Thickness.

5.3 Correlation

The correlation analysis revealed that many of the variables were significantly correlated (Table 5.6). All variables were significantly correlated with elevation. PISR and TPI were moderately positively correlated with elevation, presumably because of lesser shading and smaller loss of energy during passage through the atmosphere for the former, and a simple relationship with the higher and lower points of the landscape for the latter. The vegetation variables (canopy openness, vegetation height, and organic mat thickness) are all moderately correlated (positively for canopy openness, negatively for vegetation height and organic mat thickness) with elevation. This can be explained by the general change in vegetation characteristics with elevation. Tundra (shorter vegetation with no canopy cover and a thin organic mat) is commonly found at higher elevations while forest or shrub vegetation (with taller vegetation often with denser canopy closure and thicker organic mats) tend to be found at lower elevations. NDVI is weakly negatively correlated with elevation. Because it is a measure of vegetation greenness and vegetation cover appears at all elevations in the study areas, NDVI is not necessarily directly linked to elevation. The weak correlations between NDVI and canopy openness (negative) and vegetation height (positive) are likely a result of non-linear relationships between NDVI and the other vegetation variables.

Table 5.6: Correlation coefficients between all variables.
Significant values in bold (α=0.05).

	Elevation	PISR	TPI	NDVI	Canopy Openness	Vegetation Height	Organic Mat Thickness
Elevation	—	**0.57**	**0.38**	**-0.17**	**0.43**	**-0.52**	**-0.48**
PISR	**0.57**	—	0.03	0.08	0.01	-0.10	**-0.37**
TPI	**0.38**	0.03	—	-0.13	**0.28**	**-0.40**	**-0.50**
NDVI	**-0.17**	0.08	-0.13	—	**-0.23**	**0.26**	0.07
Canopy Openness	**0.43**	0.01	**0.28**	**-0.23**	—	**-0.87**	0.02
Vegetation height	**-0.52**	-0.10	**-0.40**	**0.26**	**-0.87**	—	0.11
Organic Mat Thickness	**-0.48**	**-0.37**	**-0.50**	0.07	0.02	0.11	—

Organic mat thickness and PISR are moderately negatively correlated and organic mat thickness, vegetation height, and canopy openness are all weakly to moderately correlated with TPI (negatively for organic mat thickness and vegetation height and positively for canopy openness).

The insignificant correlation between organic mat thickness and the other vegetation variables is not surprising. Thin organic mats occur in broadleaf forests with closed canopies as well as in tundra areas with low vegetation and no canopy closure. Similarly, thin organic mats can appear in areas with dense green vegetation (high NDVI values) as well as in areas with very little vegetation (low NDVI values). In this case, non-linear relationships between organic mat thickness and the vegetation variables may exist which would explain the low correlations.

5.4 Decision Trees

The results of the CART analysis for each of the five tree models are described in the following sections. Table 5.7 shows a comparison of the resulting accuracies, improvements, and impurities of the end nodes of the trees created using all sites. Table 4.4 (p. 46 above) identifies the input variables used for each tree. All accuracies in this section are producer's accuracies. For user's accuracies, see Appendix E.

Table 5.7: Comparison of accuracies, impurities, and improvements of the tree models created using data from all sites.

	Training Data			**Testing Data**		
Tree Model	Overall Accuracy	Impurity	Overall Improvement	Overall Accuracy	Impurity	Overall Improvement
DEM-Based tree	88.2%	0.199	0.263	73.5%	0.298	0.174
Field-Based tree	84.3%	0.234	0.228	70.6%	0.371	0.101
Field and DEM-Based tree	89.2%	0.175	0.287	76.5%	0.330	0.142
Imagery-Based tree	74.5%	0.327	0.136	55.9%	0.446	0.027
Imagery and DEM-Based tree	90.2%	0.164	0.298	64.7%	0.388	0.084

5.4.1 Selection Of Test Sample

The regression trees created following removal of the three different test sample groups varied in their structures because of differences in the data for the training groups; however, the overall accuracy of the trees remained similar (Appendix D). The test sample selected (Sample 3) was the one that produced trees that most resembled the trees created using all study sites in the training set.

5.4.2 DEM-Based

5.4.2.1 All Sites

The DEM-Based tree model used the three DEM-derived variables: elevation, PISR, and TPI. The model in which all sites were included made use of all three variables with PISR being the primary split (Figure 5.2). End nodes classified as permafrost probable occurred in five different cases: areas with low to moderately high insolation at high elevations (PISR ≤ 1406.9 MJ/m^2; elevation > 1453.5 m); areas with low insolation at low to moderately high elevations and a tendency toward concave curvature (PISR ≤ 1330.6 MJ/m^2; elevation ≤ 1453.5 m; TPI ≤ 13); areas with moderately high insolation at low to moderately high elevations (PISR 1399.6 - 1406.9 MJ/m^2; elevation ≤ 1453.5 m); areas with low insolation at moderate elevations and with a tendency toward convex curvature (PISR ≤ 1330.6 MJ/m^2; elevation 1149.5 - 1453.5 m; TPI > 13); and in areas of moderate insolation at low to moderately high elevations and with a tendency toward convex curvature (PISR 1330.6 - 1399.6 MJ/m^2; elevation ≤ 1453.5 m; TPI > 57). Permafrost improbable nodes also occurred in three difference cases: areas with high insolation (PISR > 1406.9 MJ/m^2); the areas with low insolation at low to moderate elevations and with a tendency toward convex curvature (PISR ≤ 1330.6 MJ/m^2; elevation ≤ 1149.5 m; TPI > 13); and in areas with moderate insolation at low to moderately high elevations and with a tendency toward concave curvature (PISR 1330.6 - 1399.6 MJ/m^2; elevation $\leq$ 1453.5 m; TPI $\leq$ 57).

A confusion matrix showing the correctly classified and misclassified sites for the DEM-Based tree model is shown in Table 5.8. The overall accuracy of the training data was 88.2% and the overall accuracy of the test data was 73.5%. The overall improvement of the training data was 0.263 and the final impurity was 0.199 (see p. 58, Table 5.7). This model had the lowest impurity of the test dataset (0.298).

Table 5.8: Results of the DEM-Based tree model classification.

		Predicted		
Sample	**Observed**	No Permafrost	Permafrost	Percent Correct
Training	No Permafrost	60	5	92.3%
	Permafrost	7	30	81.1%
	Overall Percentage	65.7%	34.3%	88.2%
Testing	No Permafrost	16	5	76.2%
	Permafrost	4	9	69.2%
	Overall Percentage	58.8%	41.2%	73.5%

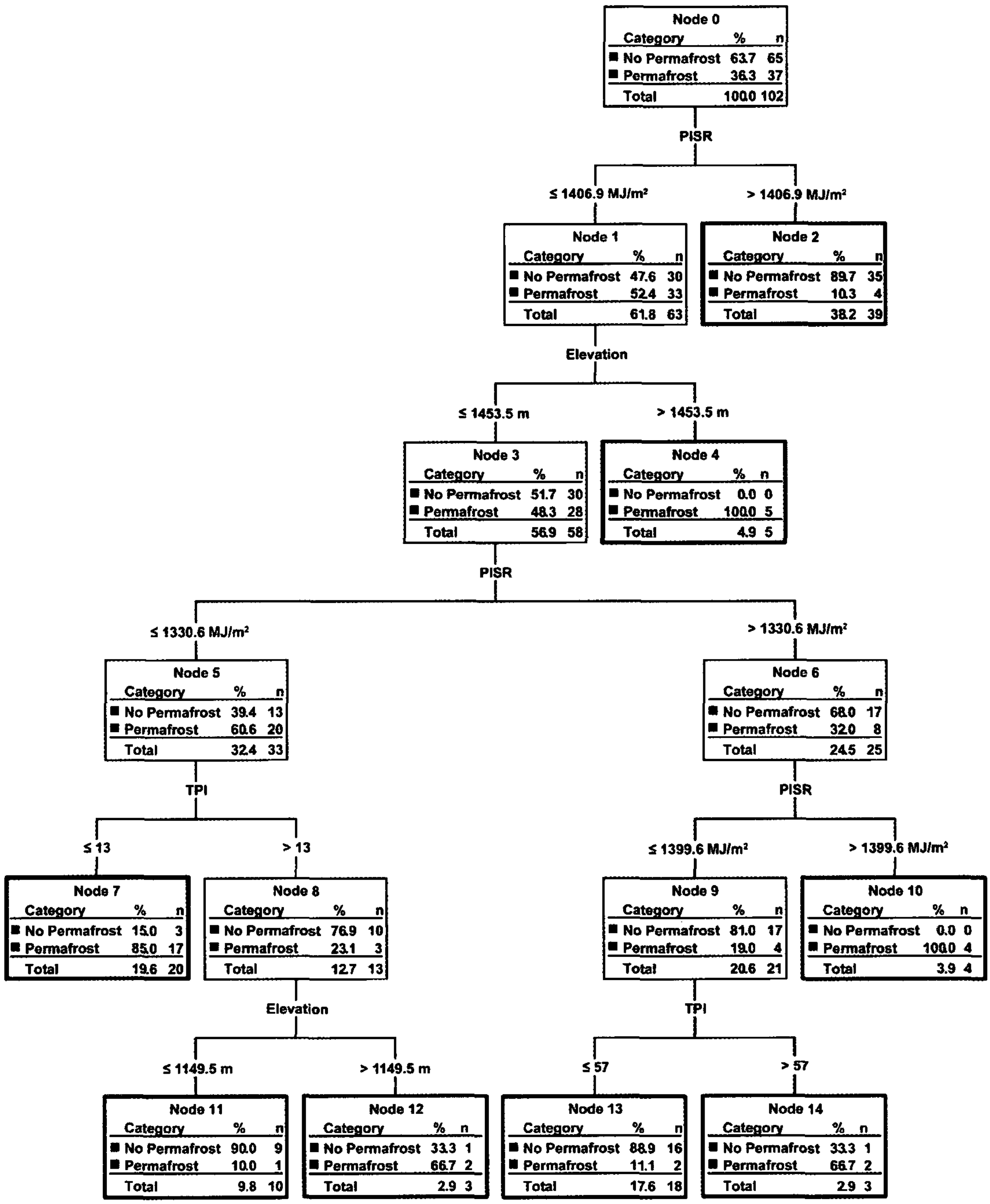


Figure 5.2: DEM-Based tree model created using all training data.
Red boxes indicate end nodes classified as permafrost improbable. Blue boxes indicate end nodes classified as permafrost probable.

5.4.2.2 Individual Study Areas

The DEM-Based tree models for the individual study areas differed from one another, but the overall accuracies were all within 9% of each other (Table 5.9).

Table 5.9: Overall accuracies, impurities, and improvements of the DEM-Based tree models for the individual sites. Partial Accuracy indicates the percentage of field sites correctly classified by the model.

	Overall Accuracy	Impurity	Overall Improvement	Partial Accuracy	
				No Permafrost	Permafrost
Dawson	93.3%	0.107	0.391	93.8%	92.9%
Faro	92.6%	0.120	0.379	92.3%	92.9%
Keno	95.5%	0.075	0.408	88.9%	100.0%
Johnson's Crossing	88.9%	0.191	0.143	95.5%	60.0%
Sa Dena Hes	86.7%	0.231	0.000	100.0%	0.0%

The Dawson model used elevation and TPI to extract four end nodes (Figure 5.3 A). The end nodes were classified as permafrost probable in three cases: at the highest elevations, at the lowest elevations, and at moderate elevations where the topography was concave (elevation > 1149.5 m; elevation ≤ 458.5 m; and elevation 458.5 - 1149.5 m and TPI ≤ 2, respectively). The sites were classified as permafrost improbable in areas at moderate elevations where the topography was convex (elevation 458.5 - 1149.5 m and TPI > 2). The overall accuracy of this model was 93.3%, the improvement was 0.391, and the final impurity was 0.107.

The Faro model, with six end nodes and using all three variables, was the most complicated base model for the individual areas (Figure 5.3 B). The three cases in which sites were classified as permafrost probable occurred at moderately low elevations with low insolation, at moderately low elevations with high insolation, and at moderately low elevations with low to moderate insolation where the topography was convex (elevation < 1199 m and PISR ≤ 1315.6 MJ/m^2; elevation < 1199 m and PISR > 1398.1 MJ/m^2; elevation < 1199 m, PISR 1315.6 - 1382.8 MJ/m^2, and TPI > 2.5). Sites were classified as permafrost improbable in areas where the elevation was moderate to high, where the elevation was moderately low with moderately high insolation, and where the elevation was moderately low with low to moderate insolation and concave topography (elevation > 1199 m; elevation ≤ 1199 m and PISR 1382.8 - 1398.1 MJ/m^2; elevation < 1199 m, PISR 1315.6 - 1382.8 MJ/m^2, and TPI ≤ 2.5). The overall accuracy was 92.6%, the improvement was 0.379, and the final impurity was 0.120. The model, however, is almost entirely counter-intuitive (e.g. no permafrost at high elevations, or permafrost at low elevations with high insolation).

The Keno DEM-Based tree model used only PISR and elevation and produced three end nodes (Figure 5.3 C). Sites were classified as permafrost probable in areas with low to moderate insolation or areas with moderate to high insolation at the highest elevations (PISR ≤ 1364.6 MJ/m^2; PISR > 1364.6 MJ/m^2 and elevation > 1456.5 m). Sites were classified as permafrost improbable in areas with moderate to high insolation at all elevations excluding only highest areas (PISR > 1364.6 MJ/m^2 and elevation ≤ 1456.5 m). This model had the highest overall accuracy (95.5%), the greatest overall improvement (0.408), and the lowest final impurity (0.075).

The Johnson's Crossing model used only one variable (PISR) to produce two end nodes (Figure 5.3 D). The sites were classified as permafrost probable in areas with low to moderate insolation, and they were classified as permafrost improbable in areas with moderate to high insolation (PISR ≤ 1374.1 MJ/m^2; PISR > 1374.1 MJ/m^2, respectively). The overall accuracy was 88.9%, the improvement was 0.143, and the final impurity was 0.191.

None of the three input variables was included in the Sa Dena Hes model. All points were therefore classified as permafrost improbable (Figure 5.3 E). This classification had the lowest accuracy (86.7%) and the highest impurity (0.231).

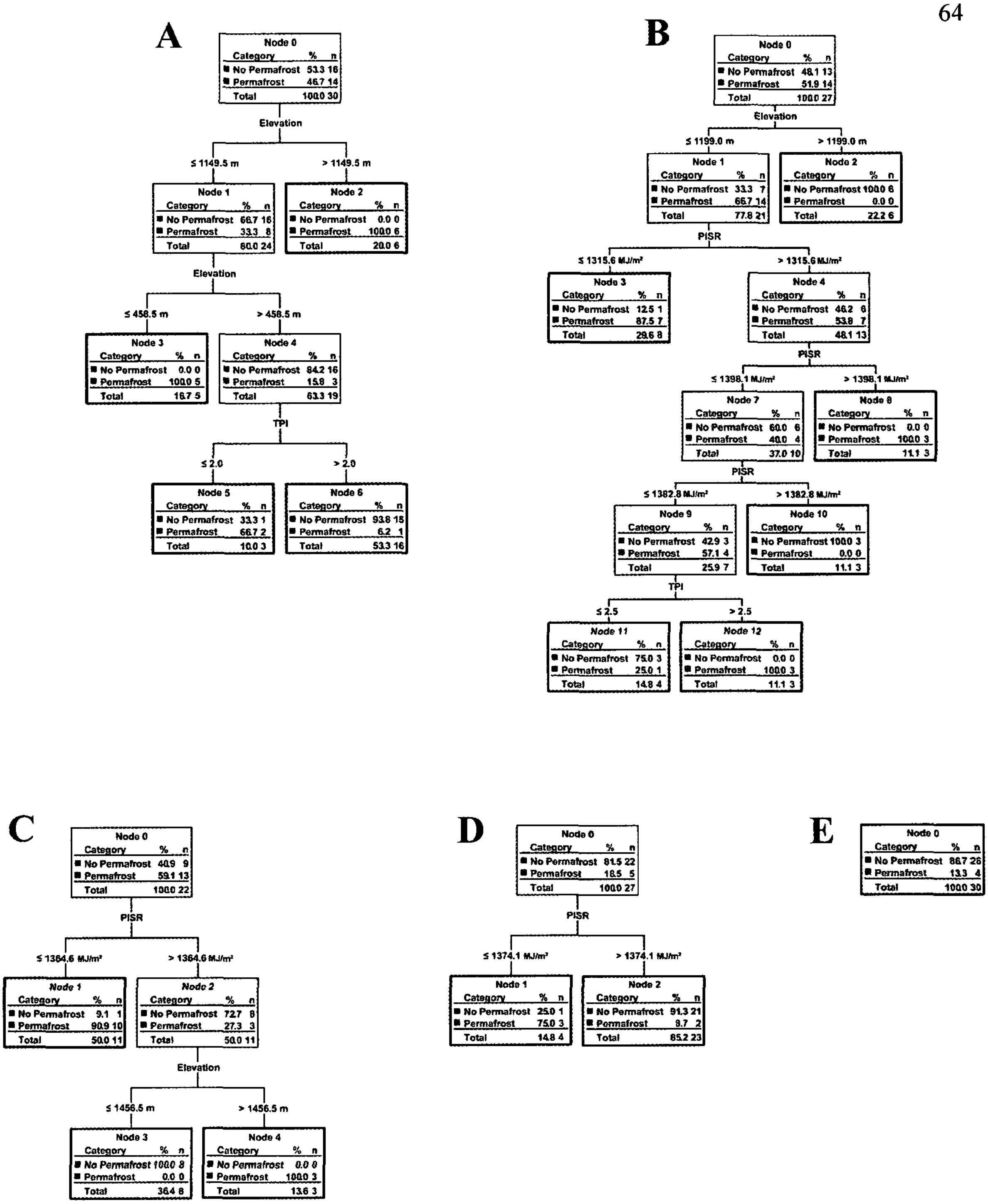


Figure 5.3: DEM-Based trees for individual study sites.
A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.

5.4.3 Field-Based

5.4.3.1 All Sites

The model for all sites containing only vegetation data collected in the field used dominant species, organic mat thickness, and vegetation height (Figure 5.4). The percentage of canopy openness was not included as it did not provide enough improvement for any splits in the model.

Nodes classified as permafrost probable occurred in four cases: where the dominant species were black spruce or bearberry (*Arctostaphylos rubra*) and the organic mat was thick (> 20.4 cm); where the dominant species were black spruce or bearberry and the organic mat was thin (≤ 10.5 cm); where the dominant species was low scrub birch (≤ 38 cm tall); and where the dominant species were subalpine fir, willow, white spruce, or alder (*Alnus crispa* spp.), in areas with all heights of vegetation excluding only those with the shortest vegetation and where the organic mat was thick (vegetation height > 38 cm; organic mat thickness > 22.15 cm). The remaining four cases were classified as permafrost improbable: in areas where the dominant species was black spruce or bearberry and the organic mat was moderately thick (10.5 - 20.4 cm); in areas where the dominant type was lichen, willow, grass, crowberry (*Empetrum nigrum*), or moss and vegetation was very short (vegetation height ≤ 38 cm); in areas where the dominant species was subalpine fir, lichen, lodgepole pine, willow, scrub birch, white spruce, grasses, trembling aspen, alder, crowberry, moss, or paper birch, including all heights of vegetation excluding the shortest, and where the organic mat was thin to moderately thick (organic mat thickness ≤ 22.15 cm); and where the dominant species was scrub birch, all heights excluding only the shortest, with a thick organic mat (vegetation height > 38 cm; organic mat thickness > 22.15 cm).

A confusion matrix showing the correctly classified and misclassified sites for the Field-Based tree model is shown in Table 5.10. The overall accuracy of the training data set was 84.3% and the test accuracy was 70.6%. The overall improvement was the second lowest at 0.228 with a final impurity of 0.234 (see p. 58, Table 5.7). The overall improvement of the test data was 0.101 and the final impurity of the test data was 0.371. The model was very successful with regard to the classification of permafrost improbable sites (98.5% correctly classified), but only 59.5% of the permafrost probable sites were correctly classified.

Table 5.10: Results of the Field-Based tree model classification.

		Predicted		
Sample	**Observed**	No Permafrost	Permafrost	Percent Correct
Training	No Permafrost	64	1	98.5%
	Permafrost	15	22	59.5%
	Overall Percentage	77.5%	22.5%	84.3%
Testing	No Permafrost	18	3	85.7%
	Permafrost	7	6	46.2%
	Overall Percentage	73.5%	26.5%	70.6%

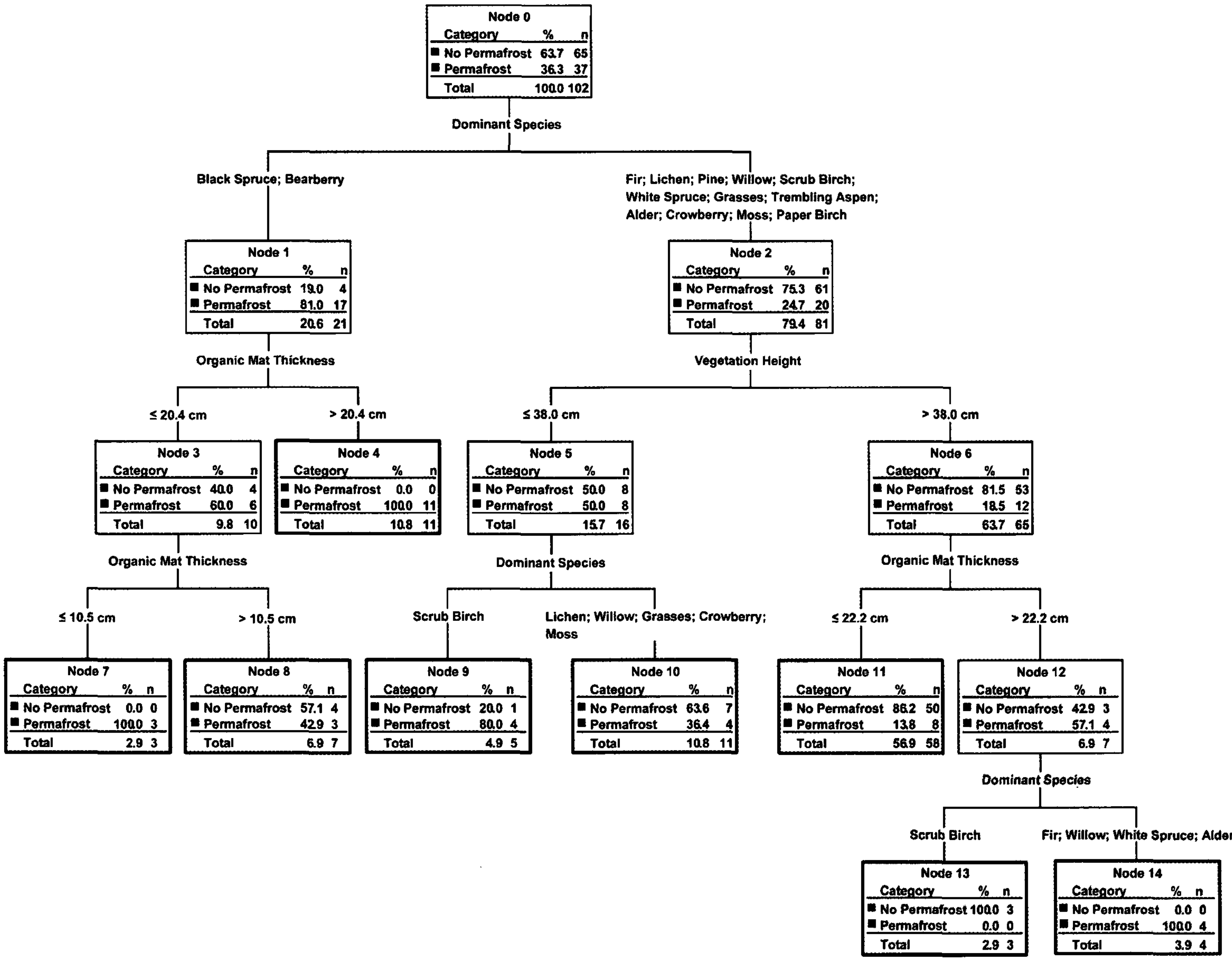


Figure 5.4: Field-Based model created using all training data.
Red boxes indicate end nodes classified as permafrost improbable. Blue boxes indicate end nodes classified as permafrost probable.

5.4.4 Field and DEM-Based

5.4.4.1 All Sites

The field data, once combined with the DEM-derived variables, produced a tree that greatly differed from the Field-Based tree (Figure 5.5). Of the seven variables input into the model, five were used (dominant species, organic mat thickness, TPI, PISR, elevation). Canopy openness and vegetation height did not provide sufficient improvement at any split to be included in the model. They also could not be used as surrogates for any other variable.

The resulting tree was complex with ten end nodes, five of which were classified as permafrost probable: in areas where the dominant species was either black spruce or bearberry and the organic mat was either very thick or very thin (organic mat thickness > 20.4 cm or ≤ 10.5 cm); where the dominant species was black spruce or bearberry, the organic mat was moderately thick, and elevation was generally lower than the surrounding area (organic mat thickness 10.5 - 20.4 cm; TPI ≤ 11); where the dominant species was subalpine fir, lichen, lodgepole pine, willow, scrub birch, white spruce, grass, trembling aspen, alder, crowberry, moss, or paper birch, where the topography was tending toward concave curvature, where the organic mat was thick, and where there was low to moderately high insolation (TPI ≤ 76.5; organic mat thickness > 22.15 cm; PISR ≤ 1406.9 MJ/m^2); or where the dominant species was either willow or scrub birch, in areas that were generally higher than surrounding areas and at low to moderate elevations (TPI > 76.5; elevation ≤ 1314 m). The remaining five end nodes were classified as permafrost improbable: where the dominant species was subalpine fir, lichen, lodgepole pine, willow, scrub birch, white spruce, grass, trembling aspen, alder, crowberry, moss, or paper birch growing on long slopes or valley bottoms where the organic mat was thin to moderately thick or growing on ridges or mountain tops where the organic mat was thick and insolation was low to moderately high (TPI ≤ 76.5 and organic mat thickness ≤ 22.15 cm; or TPI > 76.5, organic mat thickness > 22.15 cm, and PISR ≤ 1406.9 MJ/m^2); where the dominant species was lichen, grass, trembling aspen, crowberry, or paper birch and on ridges or mountain tops (TPI > 76.5); where the dominant species was black spruce or bearberry, the organic mat thickness was moderate, and elevation was higher than surrounding area (organic mat thickness 10.5 - 20.4 cm, TPI > 11); and where the dominant species was willow or scrub birch, on ridges or mountain tops at high elevations (TPI > 76.5, elevation > 1314 m).

A confusion matrix showing the correctly classified and misclassified sites for the Field and DEM-Based tree model is shown in Table 5.11. The addition of field data to the DEM-derived variables provided a 1% improvement in the overall accuracy (89.2% accuracy for the Field and DEM-Based model) (see p. 58, Table 5.7). This model also had the highest overall accuracy with the test data (76.5%) which is 3% better than the DEM-Based model. The overall improvement with the model was also higher than the DEM-Based model (0.287); however, the test data showed a lower overall improvement than the DEM-Based model test data (0.142), surprisingly given the other better results.

Compared to the DEM-Based tree (see p. 60, Table 5.8), the Field and DEM-Based model was better at predicting non-permafrost sites (98.5% vs. 92.3% for training and 95.2% vs. 76.2% for testing), but worse for permafrost sites (73.0% vs. 81.1% for training and 46.2% vs. 69.2% for testing).

Table 5.11: Results of the Field and DEM-Based tree model classification.

		Predicted		
Sample	**Observed**	No Permafrost	Permafrost	Percent Correct
Training	No Permafrost	64	1	98.5%
	Permafrost	10	27	73.0%
	Overall Percentage	72.5%	27.5%	89.2%
Testing	No Permafrost	20	1	95.2%
	Permafrost	7	6	46.2%
	Overall Percentage	79.4%	20.6%	76.5%

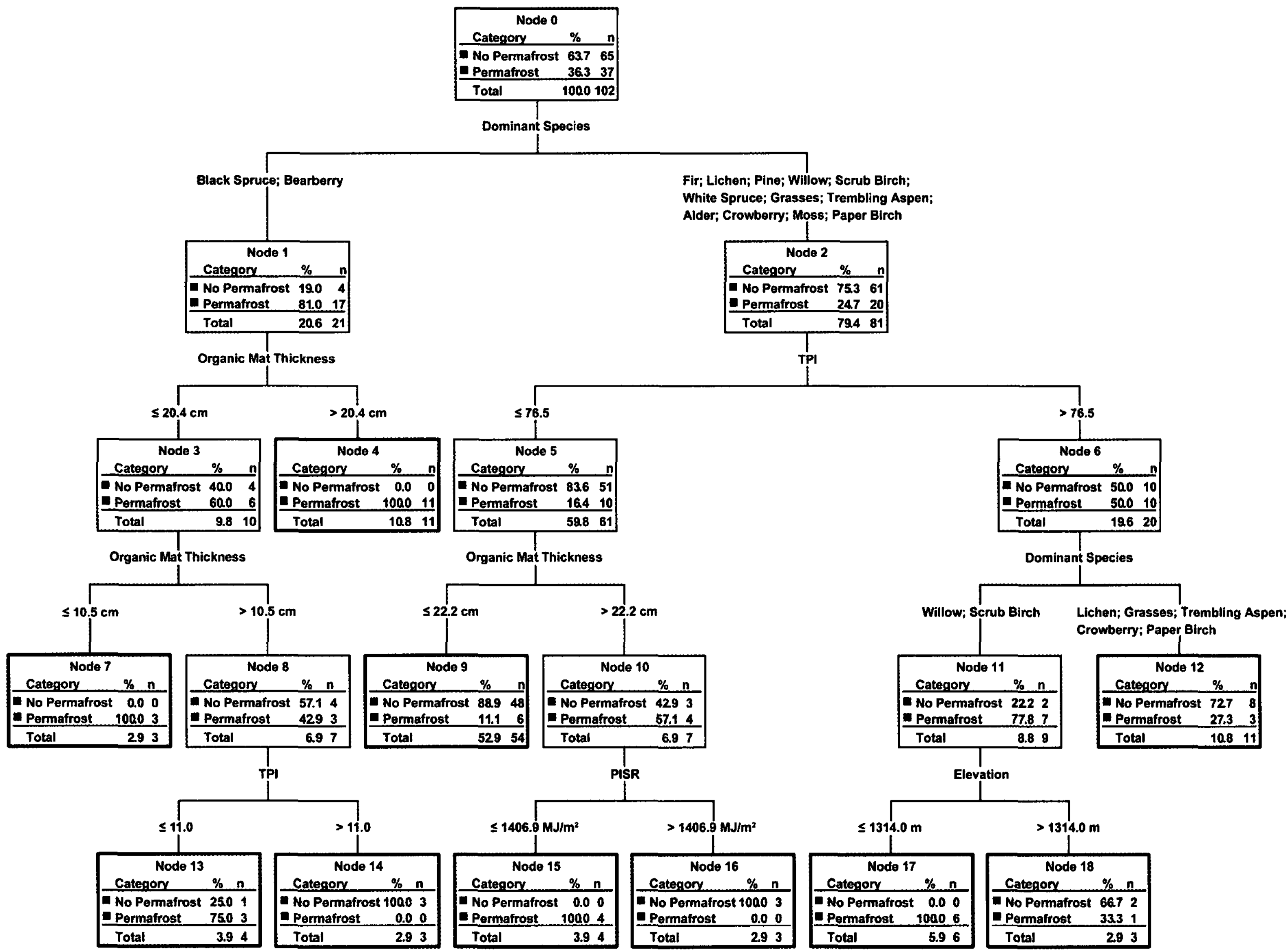


Figure 5.5: Field and DEM-Based tree model created using all training data.
Red boxes indicate end nodes classified as permafrost improbable. Blue boxes indicate end nodes classified as permafrost probable.

5.4.4.2 Field-Based and Field and DEM-Based - Individual Study Areas

The models for each individual study area did not differ between the model that included the field data only and the model that included the field data combined with the DEM-derived variables, except for Dawson. The overall accuracies ranged from 85.2% to 100%. For comparison purposes, the overall accuracies, improvements, and final impurities are displayed in Table 5.12.

Table 5.12: Overall accuracies, impurities, and improvements of the Field-Based models for the individual sites. Partial Accuracy indicates the percentage of field sites correctly classified by the model.

	Overall Accuracy	Impurity	Overall Improvement	Partial Accuracy	
				No Permafrost	Permafrost
Dawson (Field-Based)	93.3%	0.100	0.398	100.0%	85.7%
Dawson (Field and DEM-Based)	100.0%	0.000	0.498	100.0%	100.0%
Faro	85.2%	0.299	0.200	84.6%	78.6%
Keno	86.4%	0.229	0.255	88.9%	84.6%
Johnson's Crossing	92.6%	0.136	0.166	100.0%	60.0%
Sa Dena Hes	93.3%	0.114	0.117	96.2%	75.0%

The Dawson model that included only field data used the dominant species and the organic mat thickness and resulted in four end nodes (Figure 5.6 A). Sites were classified as permafrost probable in areas where the dominant species was either scrub birch or black spruce, and where the organic mat was either particularly thin or thick (organic mat thickness ≤ 10.5 cm or > 22.5 cm). Sites were classified as permafrost improbable in areas where the dominant species was white spruce, trembling aspen, grass, or paper birch, or where the dominant species was scrub birch or black spruce and the organic mat was moderately thick (organic mat thickness 10.5 - 22.5 cm). The overall accuracy of this model was 93.3%, the improvement was 0.398 and the final impurity was 0.1.

The Dawson Field and DEM-Based tree model used elevation and TPI in the place of the organic mat thickness in the field data only model (Figure 5.6 B). The two nodes classified as permafrost probable occurred when the dominant species was either scrub birch or black spruce and the elevation was either low or excluding the lowest elevations on ridges or mountain tops (elevation ≤ 705.5 m; or elevation > 705.5 m, TPI > 77). Sites were classified as permafrost improbable in areas where the dominant species was white spruce, trembling aspen, grass, or paper birch or where the dominant species was scrub birch or black spruce at all elevations excluding only the lowest on long slopes or in valleys (elevation > 705.5 m, TPI ≤ 77). The overall accuracy of this model was the highest of all models at 100% and the improvement was 0.498. The use of TPI to split Node 4 into Nodes 5 and 6 is counter-intuitive in relation to cold air drainage, but could be related to snow distribution.

The Faro models were the same and included only the dominant species (Figure 5.6 C). Sites were classified as permafrost probable in areas where the dominant species was either black spruce or white spruce. Where the dominant species was lichen, scrub birch, willow, lodgepole pine, or trembling aspen, the sites were classified as permafrost improbable. The overall accuracy of this model was the lowest at 81.5%, with an improvement of 0.2 and a final impurity of 0.299.

The tree models for Keno, Johnson's Crossing, and Sa Dena Hes also all use only the dominant species. The Keno tree classified the sites as permafrost probable when the dominant species was willow, alder, paper birch, black spruce, grass, or white spruce and as permafrost improbable when the dominant species was either scrub birch or subalpine fir (Figure 5.6 D). The overall accuracy of this model was 86.4%, the improvement was 0.255, and the final impurity was 0.229. For the Johnson's Crossing tree, the sites were classified as permafrost probable when the dominant species was lichen or black spruce and as permafrost improbable when the dominant species was willow, scrub birch, subalpine fir, lodgepole pine, white spruce, grass, or trembling aspen (Figure 5.6 E). The overall accuracy of this model was 92.6%, the improvement was 0.166, and the final impurity was 0.136. The Sa Dena Hes tree classified sites as permafrost probable when the dominant species was bearberry, black spruce, or snow willow (*Salix reticulata* ssp. *nivalis*) and as permafrost improbable when the dominant species was subalpine fir, willow, white spruce, lodgepole pine, scrub birch, alder, crowberry, grass, moss, lichen, or trembling aspen (Figure 5.6 F). The overall accuracy of this model was 93.3% with an improvement of 0.117 and a final impurity of 0.114.

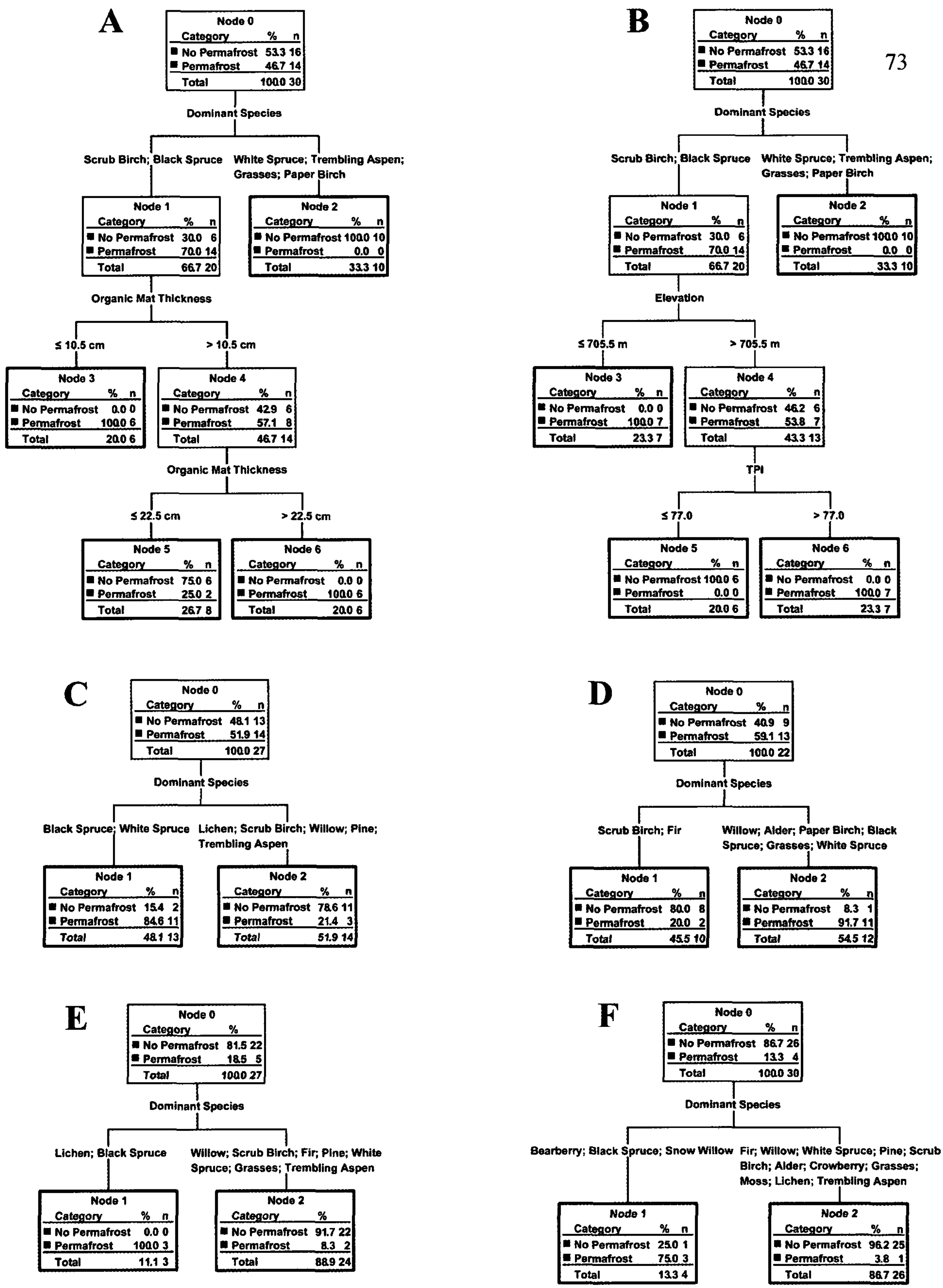


Figure 5.6: Field-Based and Field and DEM-Based models for individual study sites.
A: Dawson (Field-Based); B: Dawson (Field and DEM-Based); C: Faro; D: Keno; E: Johnson's Crossing; F: Sa Dena Hes.

5.4.5 Imagery-Based tree

5.4.5.1 All Sites

The Imagery-Based model included all three variables (the vegetation classification, NDVI, and the canopy cover classification). The tree was not very complex and had only six end nodes (Figure 5.7). The two cases in which nodes were classified as permafrost probable occurred in areas where the vegetation class was either needle-leaf or tundra and NDVI values indicated less green vegetation or unvegetated areas (≤ 0.557) and where the vegetation class was broadleaf, shrubs, or bare ground, the canopy was classified as open, and NDVI values indicated moderately green vegetation (NDVI 0.479 - 0.516). The remaining four end nodes were classified as permafrost improbable: in areas where the vegetation class was either needle-leaf or tundra and NDVI values indicated greener vegetation (NDVI > 0.557); where the vegetation class was broadleaf, shrub, or bare ground and the canopy was classified as closed; where the vegetation class was broadleaf, shrub, or bare ground, the canopy was classified as open, and NDVI values indicated greener vegetation (> 0.516); or where the vegetation class was broadleaf, shrub, or bare ground, the canopy was classified as open, and NDVI values indicated less green vegetation or unvegetated areas (≤ 0.479).

A confusion matrix showing the correctly classified and misclassified sites for the Imagery-Based model is shown in Table 5.13. This model had the lowest overall accuracy (74.5%) and the lowest accuracy of the test data (55.9%) (see p. 58, Table 5.7). The model also had the highest impurity for both the training (0.327) and testing (0.446) data and the lowest improvements for both the training (0.136) and testing (0.027) data. In the training data a very high percentage (93.8%) of the permafrost improbable sites were correctly classified. While only 40.5% of the permafrost probable sites were correctly classified. In the testing data, 76.2% of the permafrost improbable sites were correctly classified, but the vast majority of the permafrost probable sites were misclassified (only 23.1% were correct).

Table 5.13: Results of the Imagery-Based tree model classification.

Sample	Observed	Predicted: No Permafrost	Predicted: Permafrost	Predicted: Percent Correct
Training	No Permafrost	61	4	93.8%
	Permafrost	22	15	40.5%
	Overall Percentage	81.4%	18.6%	74.5%
Testing	No Permafrost	16	5	76.2%
	Permafrost	10	3	23.1%
	Overall Percentage	76.5%	23.5%	55.9%

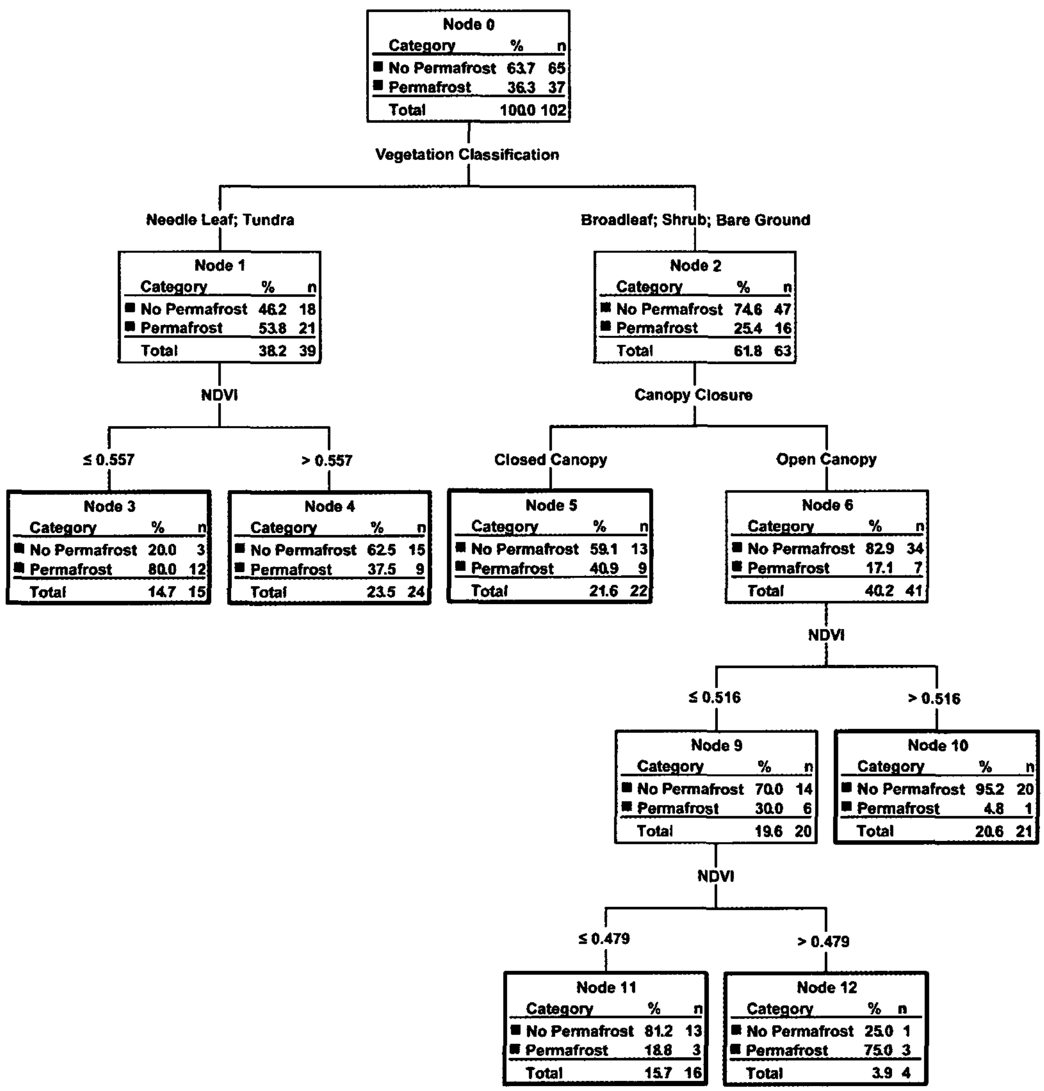


Figure 5.7: Imagery-Based tree model created using all training data.
Red boxes indicate end nodes classified as permafrost improbable. Blue boxes indicate end nodes classified as permafrost probable.

5.4.5.2 Individual Study Areas

The Imagery-Based models for each individual area had lower overall accuracies than the DEM-Based models and the models that included field data. The overall accuracies, improvements, and final impurities for the Imagery-Based models are shown in Table 5.14.

Table 5.14: Overall accuracies, impurities, and improvements of the Imagery-Based models for the individual sites. Partial Accuracy indicates the percentage of field sites correctly classified by the model.

	Overall Accuracy	Impurity	Overall Improvement	Partial Accuracy	
				No Permafrost	Permafrost
Dawson	63.3%	0.434	0.064	37.5%	92.9%
Faro	74.1%	0.376	0.123	84.6%	64.3%
Keno	77.3%	0.334	0.150	55.6%	92.3%
Johnson's Crossing	81.5%	0.302	0.000	100.0%	0.0%
Sa Dena Hes	86.7%	0.231	0.000	100.0%	0.0%

The Dawson tree model used only NDVI to produce two end nodes (Figure 5.8 A). Sites were classified as permafrost probable when NDVI values indicated less green vegetation and as permafrost improbable where NDVI values indicated very green vegetation (NDVI $\leq$ 0.670; NDVI > 0.670, repsectively). This tree produced the lowest overall accuracy of all individual site models (63.3%) with an improvement of 0.064 and the highest final impurity of 0.434.

The Faro model also only used NDVI to produce two end nodes (Figure 5.8 B). In this model the sites were classified as permafrost probable where NDVI values indicated less green or no vegetation and as permafrost improbable where NDVI values indicated greener vegetation (NDVI $\leq$ 0.547; NDVI > 0.547, respectively). The overall accuracy was 74.1%, the improvement was 0.123, and the final impurity was 0.376.

The Keno tree model was the most complicated of the Imagery-Based tree models. This tree included the vegetation class and NDVI to produce three end nodes (Figure 5.8 C). The sites were classified as permafrost probable in areas where the vegetation class was shrub, needle-leaf, or tundra or where the vegetation class was bare ground or broadleaf and the NDVI values indicated less green or no vegetation (NDVI $\leq$ 0.411). The accuracy of this model was 77.3%, the improvement was 0.150, and the final impurity was 0.334.

The Johnson's Crossing and the Sa Dena Hes models did not have any splits, so all sites were classified as permafrost improbable (Figure 5.8 D and E, respectively). The overall accuracy of the raw data for Johnson's Crossing was 81.5% and the impurity was 0.302. For Sa Dena Hes, the accuracy was 86.7% and the impurity was 0.231.

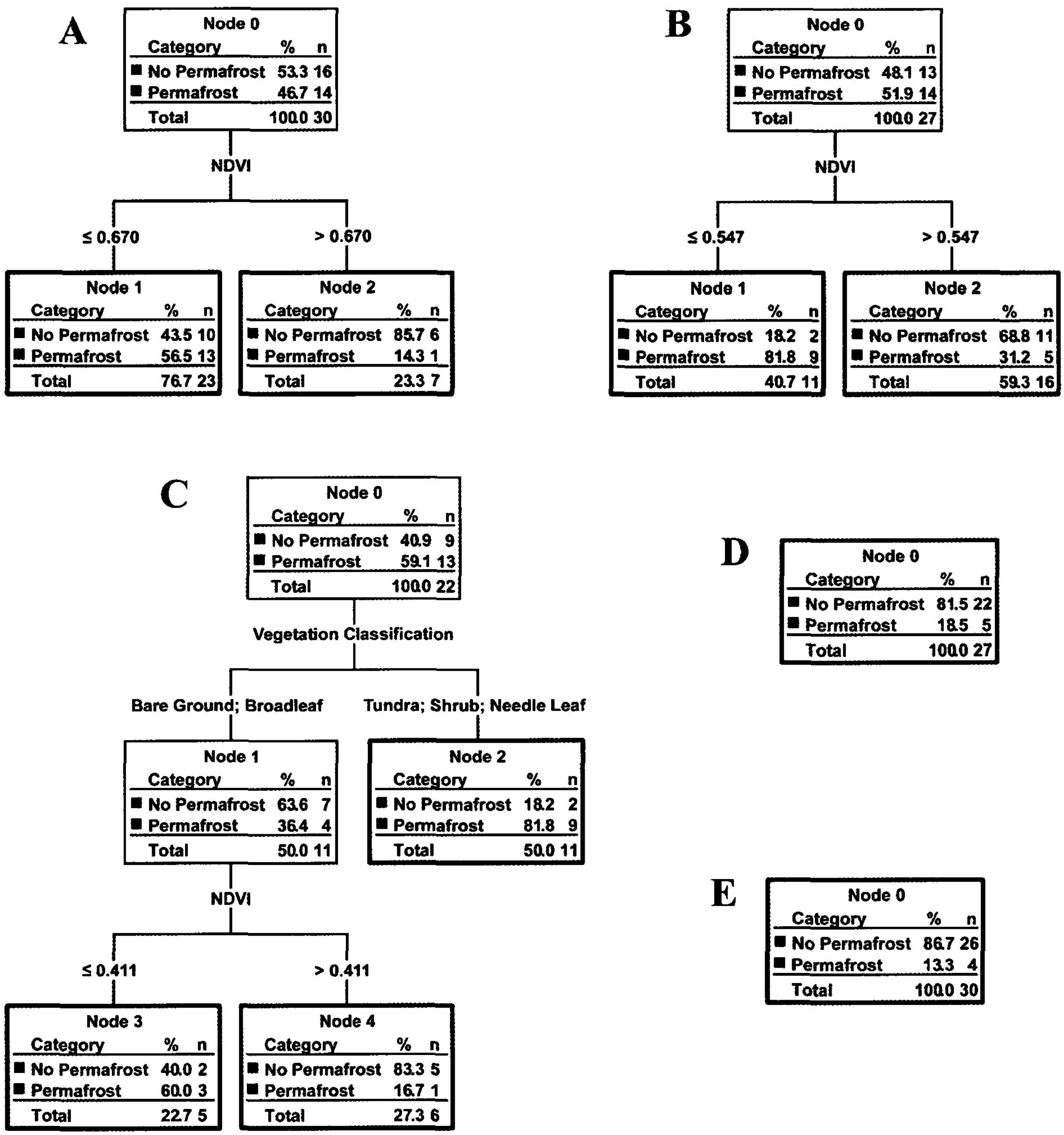


Figure 5.8: Imagery-Based models for individual study sites.
A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.

5.4.6 Imagery and DEM-Based

5.4.6.1 All Sites

The combination of satellite imagery-derived variables and DEM-derived variables created a complex model (Figure 5.9) that used four of the six input variables: PISR, NDVI, the vegetation classification, and TPI.

The model produced eleven end nodes, six of which were classified as permafrost probable: in areas where the site had low to moderately high insolation, the NDVI values indicated less green or no vegetation, and the vegetation class was either needle-leaf or tundra (PISR ≤ 1406.9 MJ/m^2; NDVI ≤ 0.553); in areas where the site had low to moderately high insolation, NDVI values indicated moderately green vegetation, the vegetation class was either needle-leaf or tundra, and the site was located on a ridge or mountain top (PISR ≤ 1406.9 MJ/m^2; NDVI 0.553 - 0.669; TPI > 68.5); in areas with low insolation where NDVI values indicated less green or no vegetation, the vegetation class was broadleaf, shrub, or bare ground, and on long slopes or in valleys (PISR ≤ 1323 MJ/m^2; NDVI ≤ 0.669; TPI ≤ 14.5); in areas with all levels of insolation excluding only the lowest and highest values where NDVI values indicated less green or no vegetation, where the vegetation class was broadleaf, shrub, or bare ground, and on ridges or mountain tops (PISR 1323 - 1406.9 MJ/m^2; NDVI ≤ 0.669; TPI > 27.5); in areas with very low insolation, where the NDVI values indicated moderately green vegetation, where the vegetation class was either needle-leaf or tundra, and where the topography formed long slopes or valleys (PISR ≤ 1258.1 MJ/m^2; NDVI 0.553 - 0.669; TPI ≤ 68.5); and in areas with very low insolation where NDVI values indicated moderately green vegetation, where the vegetation class was broadleaf, shrub or bare ground and (PISR ≤ 1323 MJ/m^2; NDVI 0.653 - 0.669; TPI > 14.5). The remaining five end nodes were classified as permafrost improbable: in areas with very high insolation (PISR > 1406.9 MJ/m^2); in areas with low to moderately high insolation where NDVI values indicated very green vegetation (PISR ≤ 1406.9 MJ/m^2; NDVI > 0.669); in areas with all but the highest and lowest insolation values where NDVI values indicated less green or no vegetation, the vegetation class was broadleaf, shrub, or bare ground, and on long slopes or in valleys (PISR 1323 - 1406.9 MJ/m^2; NDVI ≤ 0.669; TPI ≤ 27.5); in areas with all but the highest and lowest insolation values where NDVI values indicated moderately green vegetation, where the vegetation class was either needle-leaf or tundra, and on long slopes or in valleys (PISR 1258.1 - 1406.9 MJ/m^2; NDVI 0.553 - 0.669; TPI ≤ 68.5); and where insolation was low, NDVI values indicated less green or no vegetation, the vegetation class was broadleaf, shrub, or bare ground, and on mountain tops or ridges (PISR ≤ 1323 MJ/m^2; NDVI ≤ 0.653; TPI > 14.5).

A confusion matrix showing the correctly classified and misclassified sites for the Imagery and DEM-Based model is shown in Table 5.15. The overall accuracy of the training data was the highest at 90.2% which was a 2% increase over the DEM-Based model (see p. 58, Table 5.7). The test data, however, resulted in the second lowest accuracy of 64.7%. The tree is also the most complicated with the most end nodes. This model has the greatest overall improvement with the training data (0.298), but the overall improvement with the test data was the second lowest (0.084).

Table 5.15: Results of the Imagery and DEM-Based tree model classification.

		Predicted		
Sample	**Observed**	No Permafrost	Permafrost	Percent Correct
Training	No Permafrost	60	5	92.3%
	Permafrost	5	32	86.5%
	Overall Percentage	63.7%	36.3%	90.2%
Testing	No Permafrost	14	7	66.7%
	Permafrost	5	8	61.5%
	Overall Percentage	55.9%	44.1%	64.7%

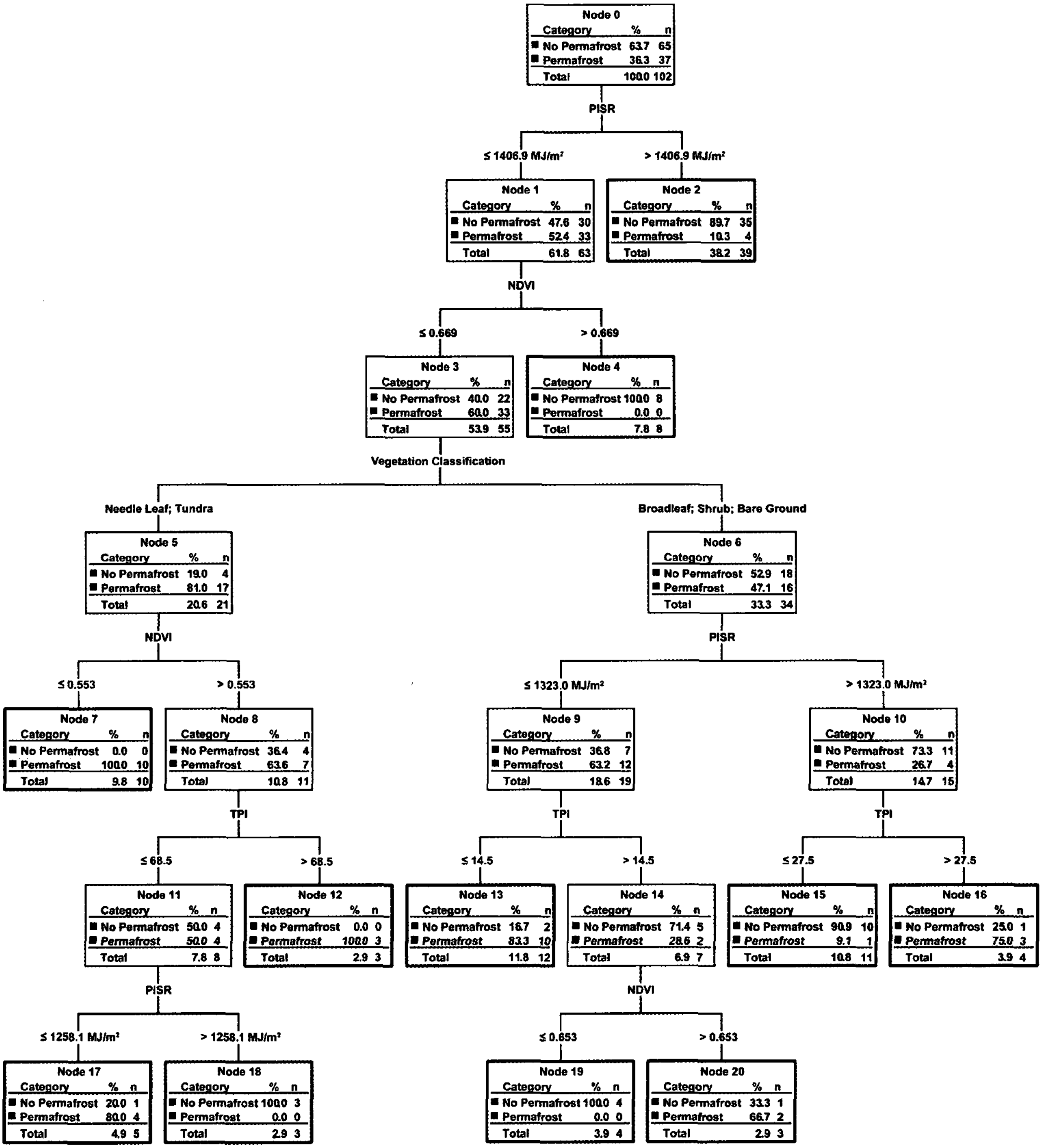


Figure 5.9: Imagery and DEM-Based model created using all training data.
Red boxes indicate end nodes classified as permafrost improbable. Blue boxes indicate end nodes classified as permafrost probable.

5.4.6.2 Individual Study Areas

The Imagery and DEM-Based tree models were identical to the DEM-Based models for three of the study areas: Dawson, Keno, and Sa Dena Hes (Figure 5.10 A, C, E, respectively). Unique models were only created for Faro and Johnson's Crossing. For comparison, all models are shown in Figure 5.10 and the descriptions from section 5.4.2.2 apply. The overall accuracies, improvements, and final impurities for the Imagery and DEM-Based models are shown in Table 5.16.

Table 5.16: Overall accuracies, impurities, and improvements of the Imagery and DEM-Based models for the individual sites. Partial Accuracy indicates the percentage of field sites correctly classified by the model.

	Overall Accuracy	Impurity	Overall Improvement	Partial Accuracy	
				No Permafrost	Permafrost
Dawson	93.3%	0.107	0.391	93.8%	87.5%
Faro	85.2%	0.165	0.335	69.2%	100.0%
Keno	95.5%	0.075	0.408	88.9%	100.0%
Johnson's Crossing	92.6%	0.105	0.197	90.9%	100.0%
Sa Dena Hes	86.7%	0.231	0.000	100.0%	0.0%

The Faro tree used elevation, NDVI, and TPI to produce three splits and four end nodes (Figure 5.10 B). Sites were classified as permafrost probable in two cases: where the elevation was moderately low and where NDVI values indicated less green or no vegetation (elevation ≤ 1199 m; NDVI ≤ 0.547); or where the elevation was moderately low, NDVI values indicated very green vegetation, and surrounding elevations tended to be lower (elevation ≤ 1199 m; NDVI > 0.547; TPI > -23.5). Sites were classified as permafrost improbable in two cases: where the elevation was moderate to high (elevation > 1199 m); or where the elevation was moderately low, NDVI values indicated very green vegetation, and surrounding elevations were higher (elevation ≤ 1199 m; NDVI > 0.546; TPI ≤ -23.5). The overall accuracy of this model was 85.2%, the improvement was 0.335, and the final impurity was 0.165. As was the case for the DEM-Based model for Faro, the classification of permafrost improbable at high elevations was unexpected.

The Johnson's Crossing tree used PISR and NDVI and resulted in three end nodes (Figure 5.10 D). Sites that were classified as permafrost probable occurred in two cases: areas with low to moderate insolation (PISR $\leq$ 1374.1 MJ/m^2); or, surprisingly, areas with moderate to high insolation where NDVI values indicated very green vegetation (PISR > 1374.1 MJ/m^2; NDVI > 0.684). Sites that were classified as permafrost improbable occurred with moderate to high insolation where NDVI values indicated less green or no vegetation (PISR > 1374.1 MJ/m^2; NDVI $\leq$ 0.684). The overall accuracy of this model was 92.6%, the improvement was 0.197, and the final impurity was 0.105.

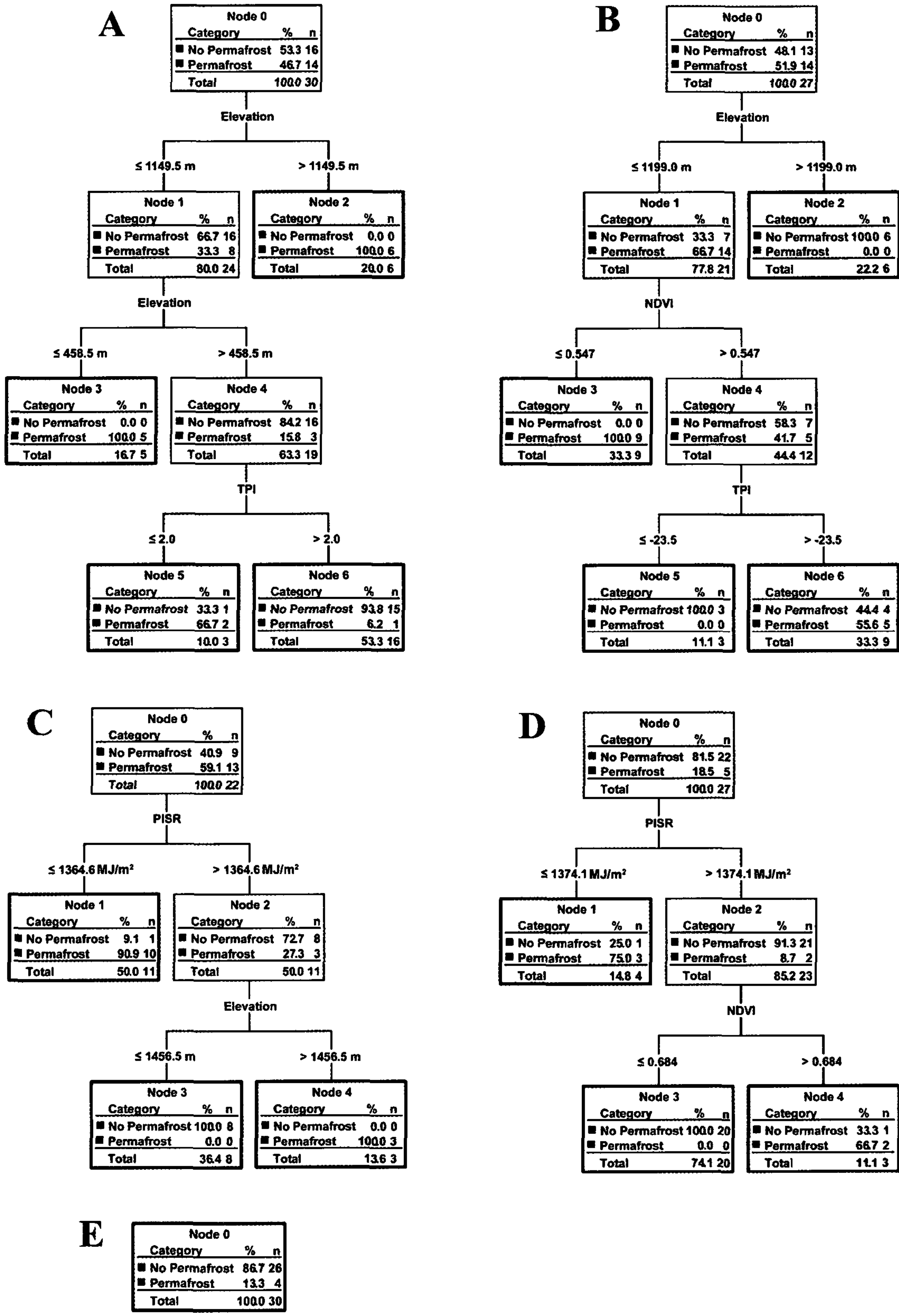


Figure 5.10: Imagery and DEM-Based models for individual study sites.
A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.

5.5 Modelled Permafrost Maps

The maps created using the DEM-Based, Imagery-Based, and Imagery and DEM-Based tree models are displayed in Figures 5.11 to 5.15. Table 5.17 shows the percentage of each study area that is classified as permafrost probable.

Table 5.17: Percentage of each study area that is classified as permafrost probable by each model in each study area. *A: Dawson; B: Faro; C: Keno; D: Johnson's Crossing; E: Sa Dena Hes.*

A	% Area Classified as Permafrost Probable
DEM-Based tree (all sites)	53.4%
DEM-Based tree (individual)	46.6%
Imagery-Based (all sites)	2.3%
Imagery-Based (individual)	46.5%
Imagery and DEM-Based (all sites)	33.7%
Imagery and DEM-Based (individual)	46.6%

B	% Area Classified as Permafrost Probable
DEM-Based tree (all sites)	39.2%
DEM-Based tree (individual)	40.8%
Imagery-Based (all sites)	33.6%
Imagery-Based (individual)	53.2%
Imagery and DEM-Based (all sites)	42.1%
Imagery and DEM-Based (individual)	55.2%

C	% Area Classified as Permafrost Probable
DEM-Based tree (all sites)	51.5%
DEM-Based tree (individual)	73.2%
Imagery-Based (all sites)	20.2%
Imagery-Based (individual)	55.9%
Imagery and DEM-Based (all sites)	46.4%
Imagery and DEM-Based (individual)	73.2%

D	% Area Classified as Permafrost Probable
DEM-Based tree (all sites)	10.6%
DEM-Based tree (individual)	30.9%
Imagery-Based (all sites)	5.5%
Imagery-Based (individual)	0.0%
Imagery and DEM-Based (all sites)	4.2%
Imagery and DEM-Based (individual)	41.0%

E	% Area Classified as Permafrost Probable
DEM-Based tree (all sites)	11.1%
DEM-Based tree (individual)	0.0%
Imagery-Based (all sites)	1.1%
Imagery-Based (individual)	0.0%
Imagery and DEM-Based (all sites)	1.6%
Imagery and DEM-Based (individual)	0.0%

5.5.1 Dawson

The modelled permafrost distribution maps for Dawson are shown in Figure 5.11. Figure 5.11 A shows the results of the DEM-Based tree model that used all data. Permafrost is probable at higher and lower elevations while it is improbable at middle elevations. Approximately 53.4% of the area has been classified as probable permafrost. Figure 5.11 B shows the map produced by the DEM-Based tree model for the Dawson area. This map suggests that permafrost is probable over 46.6% of the study area. Like Map A, Map B shows probable permafrost at lower and higher elevations while the middle elevations are classified as permafrost improbable. The main difference between the two maps is that the permafrost probable class at higher elevations is much less extensive in the map created using only Dawson data.

The result of the Imagery-Based tree model created using all study sites is shown in Figure 5.11 C. Approximately 2.3% of the area is classified as permafrost probable and this classification is mostly located at higher elevations. Figure 5.11 D is the map produced using the Imagery-Based tree model created from only sites within the Dawson study area. This model is based solely on NDVI values. The resulting 46.5% of the area that is classified as permafrost probable is at a range of elevations from the lowest to the highest with a tendency toward coverage of north-facing slopes.

The map produced from the Imagery and DEM-Based tree model for all sites is shown in Figure 5.11 E. The 33.7% of the area classified as permafrost probable shows permafrost at the lowest elevations, at the highest elevations, and on north-facing slopes. Figure 5.11 F shows the same prediction map as Figure 5.11 B because the Imagery and DEM-Based tree model for Dawson did not include any imagery-derived data.

The maps that best represent the permafrost coverage for the area based on the accuracy of the tree models are Maps B and F. Dawson is within the extensive discontinuous permafrost zone with 50% - 90% of the area underlain by permafrost (Heginbottom *et al.*, 1995). Only Map A fits into this classification.

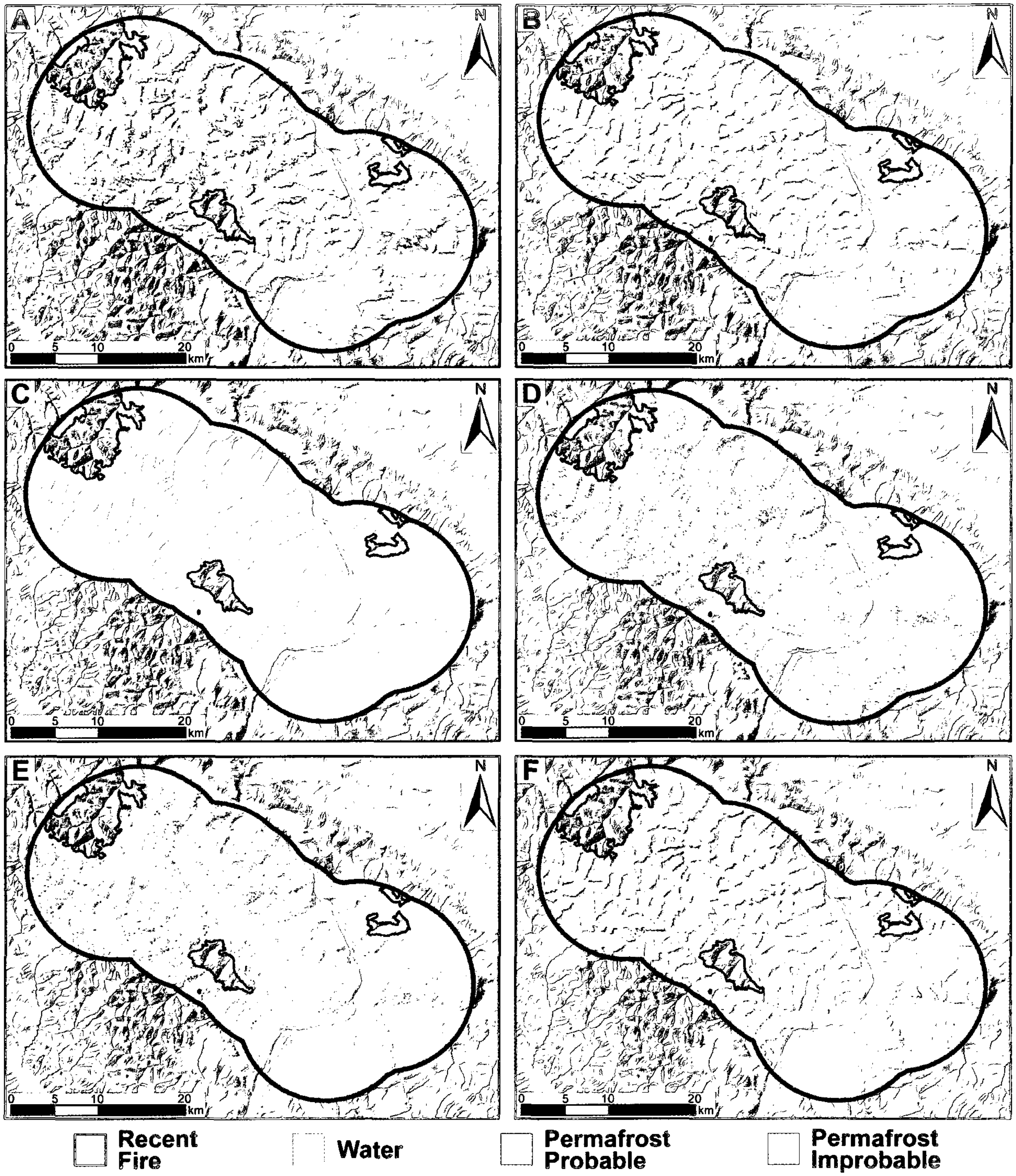


Figure 5.11: Permafrost distribution prediction maps for Dawson based on each tree model. *A: DEM-Based, all sites (Figure 5.2); B: DEM-Based, Dawson sites (Figure 5.3 A); C: Imagery-Based, all sites (Figure 5.7); D: Imagery-Based, Dawson sites (Figure 5.8 A); E: Imagery and DEM-Based, all sites (Figure 5.9); F: Imagery and DEM-Based, Dawson sites (Figure 5.10 A).*

5.5.2 Faro

Figure 5.12 shows the permafrost distribution maps produced for the Faro study area. The map produced from the base model created using all data is shown in Figure 5.12 A. The map displays probable permafrost at the lowest elevations and on north-facing slopes. The only areas at higher elevations classified as probable permafrost are on north-facing slopes. A total of 39.2% of the study area is covered by the permafrost probable class. Figure 5.12 B is the map produced from the base model with only Faro study sites. In this model, probable permafrost areas are located on both north- and south-facing slopes and are particularly concentrated along the Pelly River. The model indicates that permafrost is improbable at higher elevations. This model classifies 40.8% of the study area as permafrost probable. The CART model used for this map contained splits with predictions that are opposite to what would be expected (e.g. no permafrost at high elevations, and permafrost at low elevations with high insolation).

The Imagery-Based tree model created from all data classified only 33.6% of the study area as permafrost probable (Figure 5.12 C). On this map, the permafrost probable class is scattered throughout with concentrations at higher elevations and, interestingly, on south-facing slopes at lower elevations. Figure 5.12 D shows the map produced from the Imagery-Based tree model created using only Faro study sites. This model classified the entire mine site as probable permafrost as well as the highest elevations and scattered locations at lower elevations. 53.2% of the area is classified as probable permafrost.

The map produced using the Imagery and DEM-Based tree model created using all sites is shown in Figure 5.12 E. On this map, permafrost is classified as probable at lower elevations along the river and on north-facing slopes. The model classified 42.1% of the area as permafrost probable. Figure 5.12 F shows the map produced using the Imagery and DEM-Based tree model for only Faro sites. The areas classified as probable permafrost cover 55.2% of the area and are located throughout the mine site and at lower elevations. As for Map B, the model used for this map contained an unexpected split where sites were classified as permafrost improbable at higher elevations

The map that shows the best representation of the permafrost coverage based on the accuracy of the tree model is Map B. Faro is on the border between the sporadic discontinuous and extensive discontinuous permafrost zones (10-50% and 50-90% of the area underlain by permafrost, respectively), so approximately 50% of the terrain is expected to be underlain by permafrost (Heginbottom *et al.*, 1995). All maps come reasonably close to this, with values from 33 to 55%.

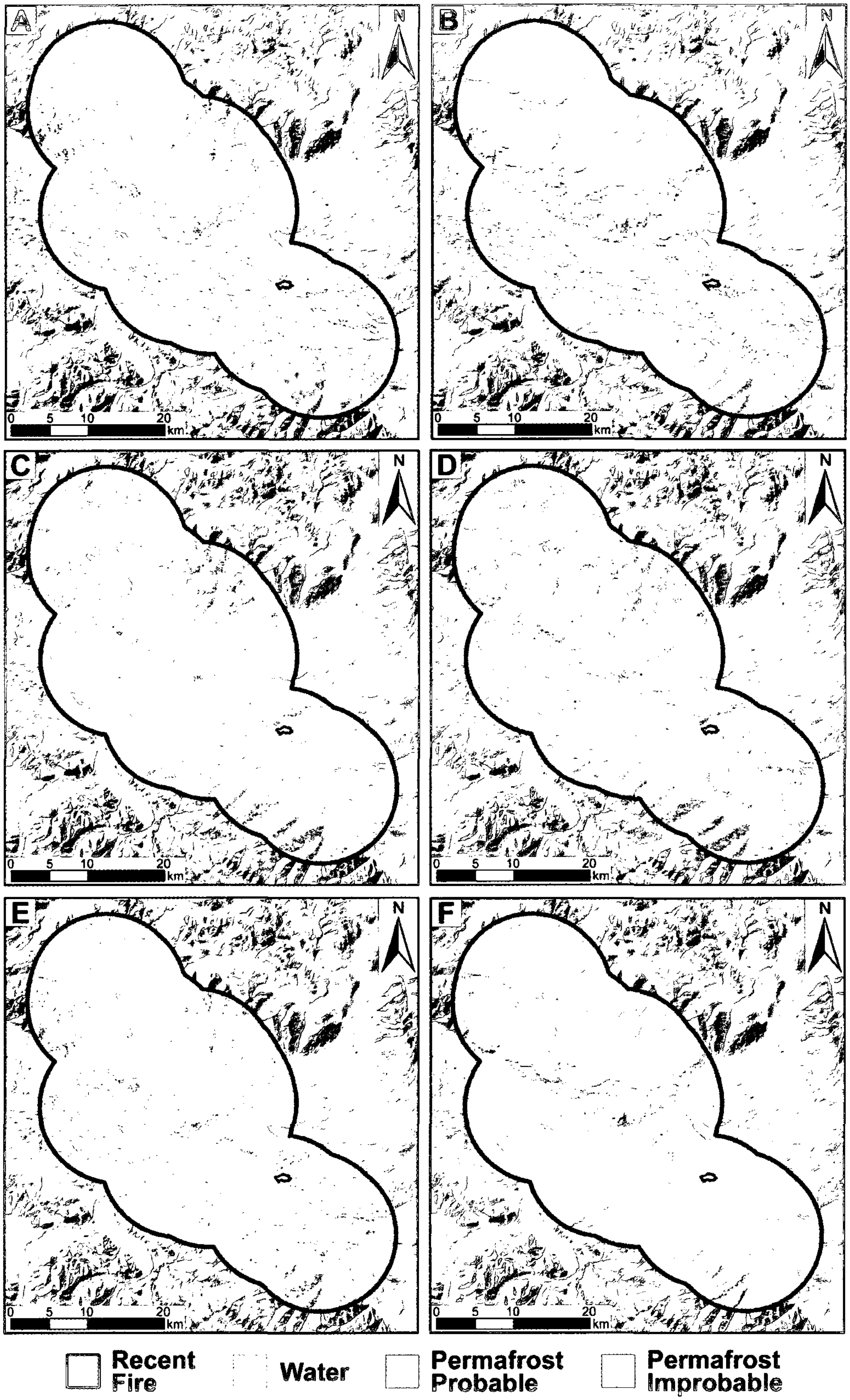


Figure 5.12: Permafrost distribution prediction maps for Faro based on each tree model.
A: DEM-Based, all sites (Figure 5.2); B: DEM-Based, Faro sites (Fgure 5.3 B); C: Imagery-Based, all sites (Figure 5.7); D: Imagery-Based, Faro sites (Figure 5.8 B); E: Imagery and DEM-Based, all sites (Figure 5.9); F: Imagery and DEM-Based, Faro sites (Figure 5.10 B).

5.5.3 Keno

The maps produced for the Keno area are shown in Figure 5.13. On the map produced from the DEM-Based model using all sites, shown in Figure 5.13 A, permafrost is classified as probable at lower elevations, and extends to higher elevations on north-facing slopes. Of the area, 51.5% was classified as probable permafrost. Figure 5.13 B shows the map created from the DEM-Based model for just the Keno sites. This map is similar to Map A, but the permafrost probable class extends higher up the valleys. The only permafrost improbable patches are at middle to higher elevations on south-facing slopes. Nearly three quarters of the study area (73.2%) was classified as probable permafrost.

Figure 5.13 C shows the map produced from the Imagery-Based tree model for all sites. This map is quite different from the other permafrost distribution maps produced for Keno. The majority of the area is classified as permafrost improbable. Patches of permafrost probable, comprising 20.2% of the study area, are scattered around lower elevations and at the tops of mountains with a tendency toward south-facing slopes. In contrast to Map C, Figure 5.13 D has the permafrost probable class scattered throughout the area with some concentration at higher elevations. This model classified 55.9% of the study area as permafrost probable.

The Imagery and DEM-Based tree model for all sites produced a similar permafrost distribution map to Map A (Figure 5.13 E). The main difference between the two is that Map E has less probable permafrost on north-facing slopes and more on ridges and mountain tops. The Imagery and DEM-Based tree model for the Keno sites contained no imagery-derived data; therefore, Map F is the same as Map B.

The best representations of permafrost coverage based on the accuracy of the tree model were Maps B and F. Keno is entirely within the extensive discontinuous permafrost zone with 50-90% of the terrain underlain by permafrost (Heginbottom *et al.*, 1995). Maps A, B, D, and F fit into this classification.

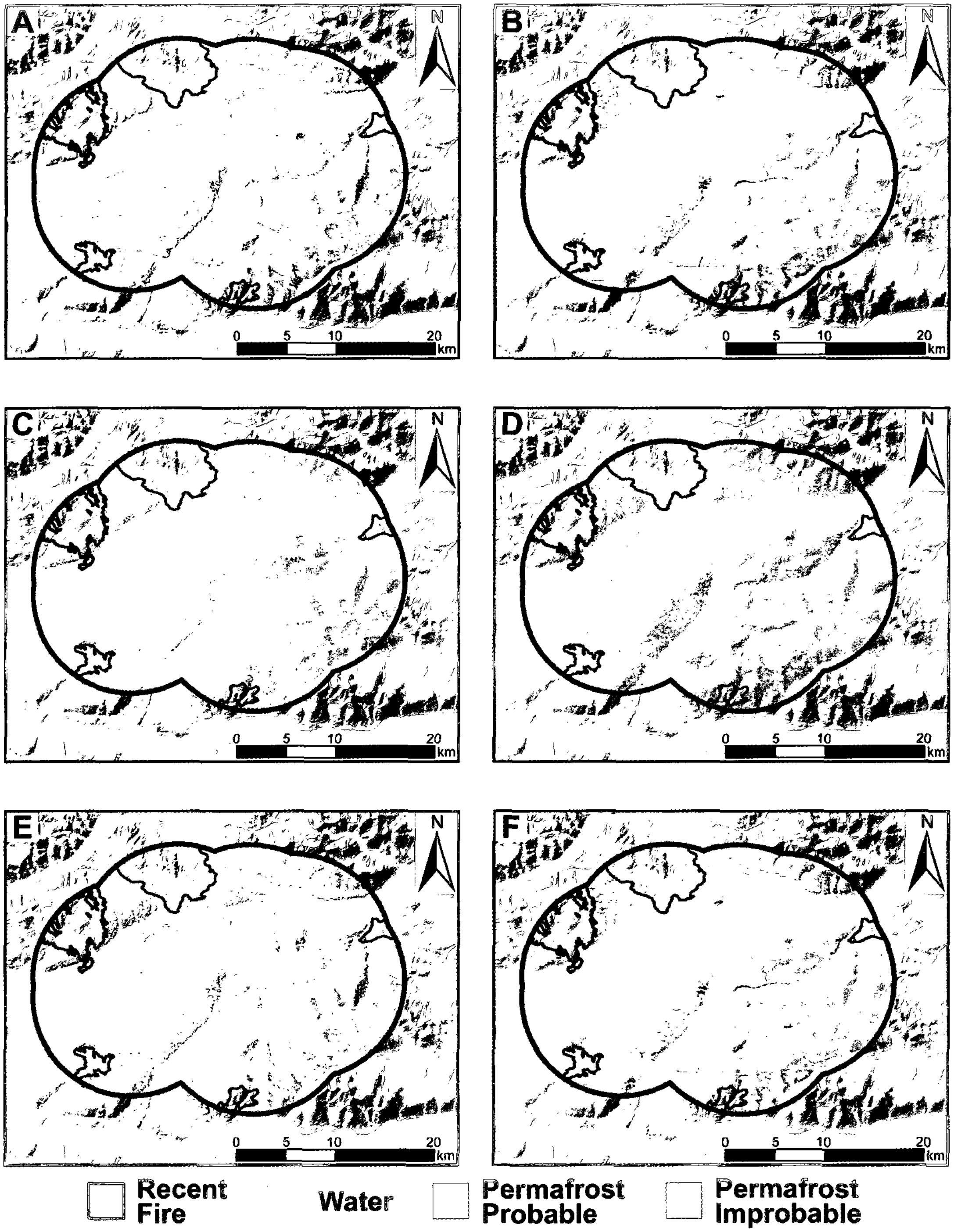


Figure 5.13: Permafrost distribution prediction maps for Keno based on each tree model. *A: DEM-Based, all sites (Figure 5.2); B: DEM-Based, Keno sites (Figure 5.3 C); C: Imagery-Based, all sites (Figure 5.7); D: Imagery-Based tree, Keno sites (Figure 5.8 C); E: Imagery and DEM-Based, all sites (Figure 5.9); F: Imagery and DEM-Based, Keno sites (Figure 5.10 C).*

5.5.4 Johnson's Crossing

Maps created for the Johnson's Crossing study area are shown in Figure 5.14. Figure 5.14 A shows the map produced from the DEM-Based tree model for all sites. This map shows few areas with probable permafrost (10.6% of the area) primarily at higher elevations and on north-facing slopes. In the map produced from the DEM-Based tree model for Johnson's Crossing sites, shown in Figure 5.14 B, the permafrost probable class, covering 30.9% of the study area, was concentrated at lower elevations on the north and south sides of the study area with some at higher elevations on north-facing slopes.

The map produced by the Imagery-Based tree model for all data is shown in Figures 5.14 C. This model classified only very small patches, covering 5.5% of the entire area, at the highest elevations on the map as probable permafrost. Figure 5.14 D shows the map produced by the Imagery-Based tree model for Johnson's Crossing sites only. The entire area was classified as permafrost improbable because the model contained no splits.

The map produced by the Imagery and DEM-Based tree model for all sites classified very few areas as permafrost probable (only 4.2% of the entire area), and only at the highest elevations (Figure 5.14 E). The map produced from the model using only Johnson's Crossing points, shown in Figure 5.14 F, is very similar to Map B with an additional permafrost probable class at middle elevations. A total of 41% of the study area was classified as probable permafrost using this model.

Map F was the best representation of permafrost coverage based on the accuracy of the tree model. Johnson's Crossing is within the sporadic discontinuous permafrost zone with 10-50% of the area underlain by permafrost (Heginbottom *et al.*, 1995). Maps A, B, and F fit into this classification.

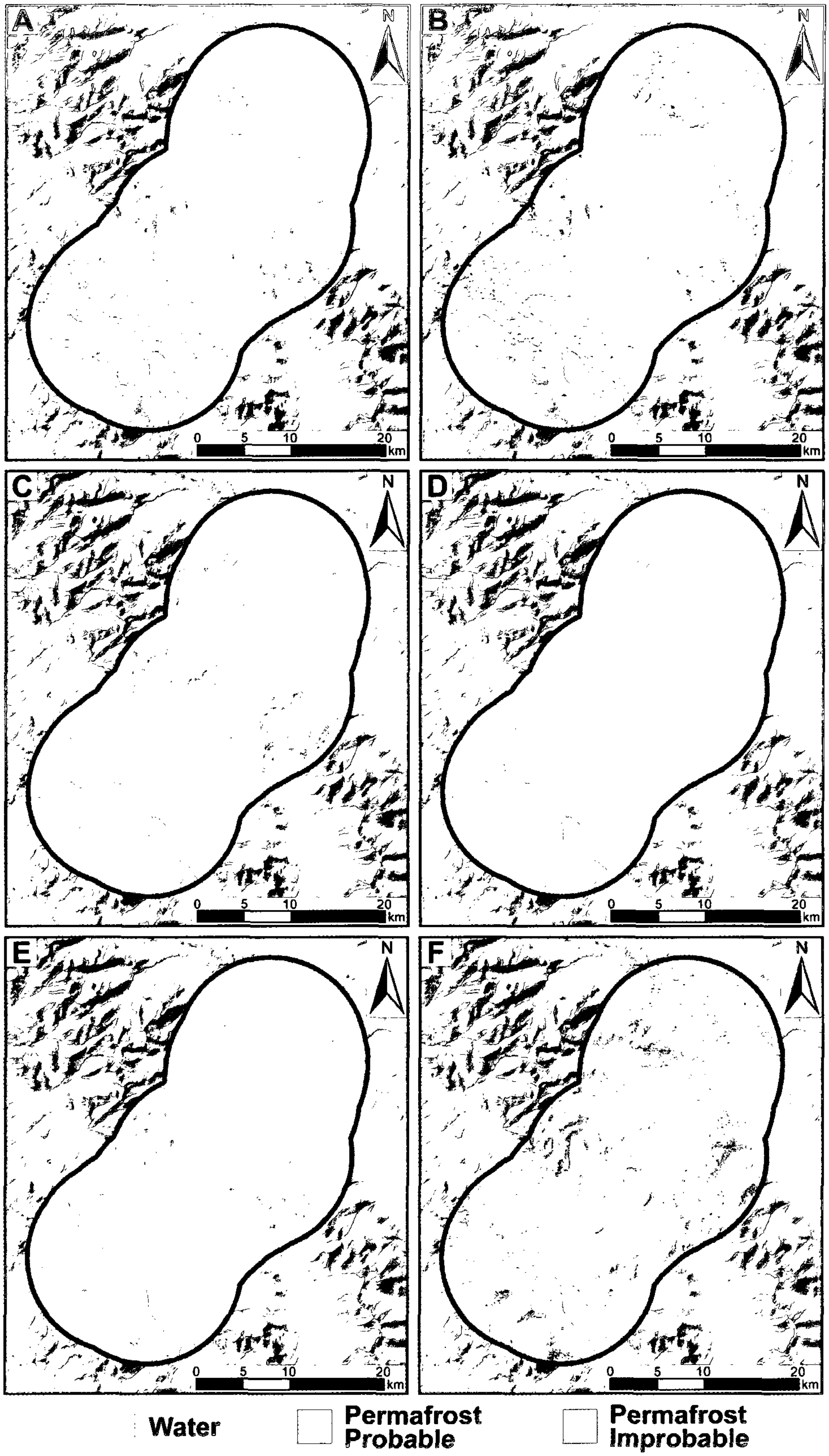


Figure 5.14: Permafrost distribution prediction maps for Johnson's Crossing based on each tree model.
A: DEM-Based, all sites (Figure 5.2); B: DEM-Based, Johnson's Crossing sites (Figure 5.3 D); C: Imagery-Based, all sites (Figure 5.7); D: Imagery-Based, Johnson's Crossing sites (Figure 5.8 D); E: Imagery and DEM-Based, all sites (Figure 5.9); F: Imagery and DEM-Based, Johnson's Crossing sites (Figure 5.10 D).

5.5.5 Sa Dena Hes

The map produced from the DEM-Based model for all sites is shown in Figure 5.15 A. The permafrost probable class covered 11.1% of the area and was distributed on north-facing slopes and in some valleys. The map using the model created solely from Sa Dena Hes sites predicted no permafrost probable areas because the tree contained no splits (Figure 5.15 B).

The map created from the Imagery-Based tree model for all sites is shown in Figure 5.15 C. Only a small portion (1.1%) of the area at the tops of mountains and in a couple of scattered areas at lower elevations was classified as permafrost probable. Like Map B, the Imagery-Based model created with only Sa Dena Hes sites classified the entire study area as permafrost improbable (Figure 5.15 D).

Figure 5.15 E shows the map produced from the Imagery and DEM-Based tree model for all sites. The distribution of probable permafrost in this map is similar to that in Map C, but there are more (smaller) patches of probable permafrost (covering 1.6% of the area). Just like the other models created with only Sa Dena Hes sites, the model contained no splits, thus all points were classified as permafrost improbable for Figure 5.15 F.

Based on the accuracy of the tree model, Map E was the best representation of permafrost distribution. However, Sa Dena Hes is within the sporadic discontinuous permafrost zone with 10-50% of the area underlain by permafrost (Heginbottom *et al.*, 1995). Only Map A fits into this classification.

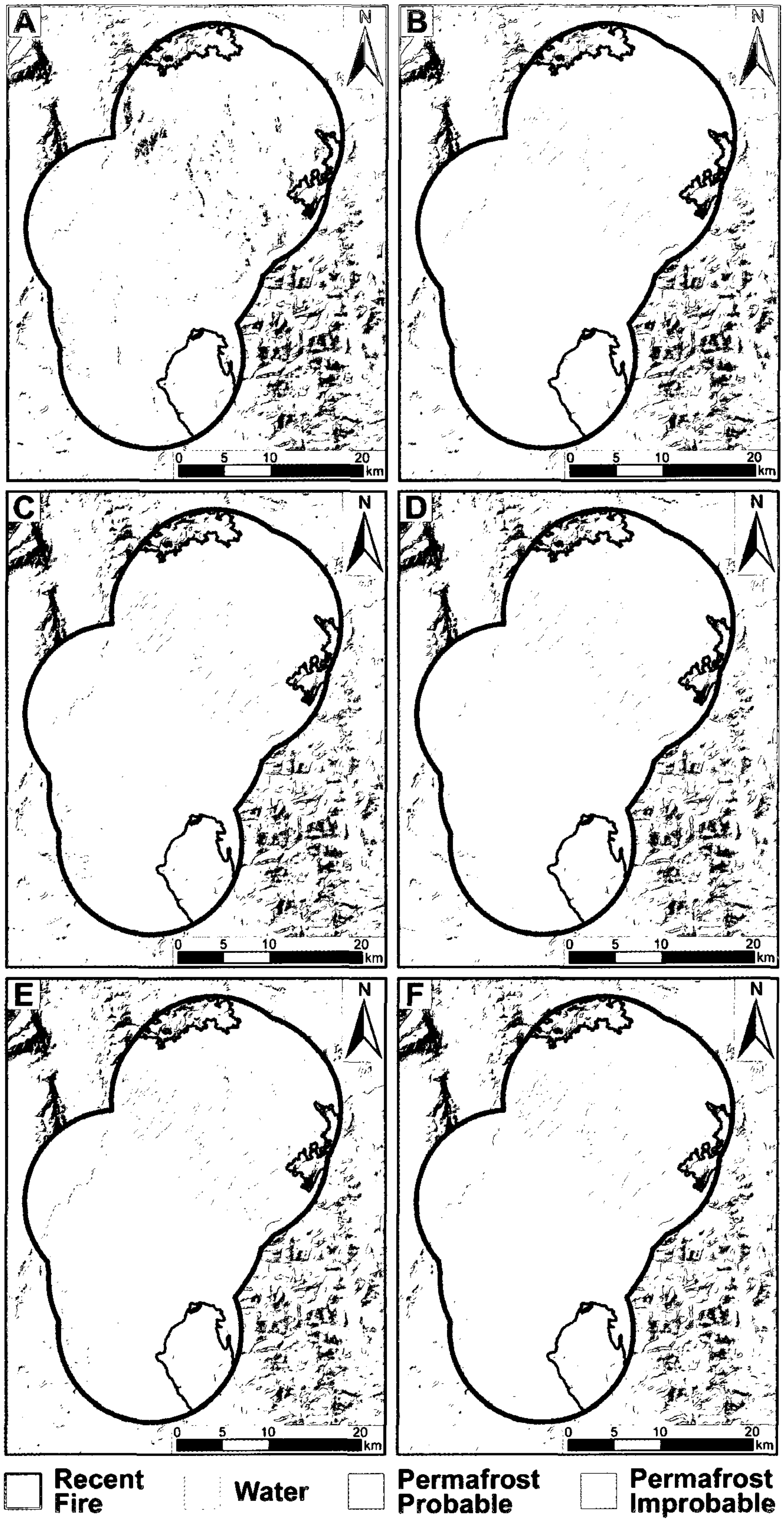


Figure 5.15: Permafrost distribution prediction maps for Sa Dena Hes based on each tree model
A: DEM-Based, all sites (Figure 5.2); B: DEM-Based, Sa Dena Hes sites (Figure 5.3 E); C: Imagery-Based, all sites (Figure 5.7); D: Imagery-Based, Sa Dena Hes sites (Figure 5.8 E); E: Imagery and DEM-Based, all sites (Figure 5.9); F: Imagery and DEM-Based, Sa Dena Hes sites (Figure 5.10 E).

6 DISCUSSION

Examination of the initial comparisons of the variables and the final decision tree models revealed a number of interesting results.

6.1 VARIABLES AND SURROGATES

6.1.1 CANOPY OPENNESS AND CANOPY CLOSURE

Only one of the ten variables used as inputs for the models was not used in any tree. The percentage of canopy openness (from the data collected in the field) was not useful in any model. A number of reasons might explain why this variable is not useful. First, based on the difference of means test, canopy openness is not significantly different between permafrost and non-permafrost sites. This variable is also significantly correlated with four other variables (elevation, TPI, NDVI, and vegetation height) that appear in the tree models. The correlations between canopy openness and elevation, and canopy openness and vegetation height are not surprising because as elevation increases, vegetation height decreases, and thus canopy openness increases as well. Moreover, all sites with tundra or shrub vegetation were labelled as 100% canopy openness; therefore, there is an inherent association between the height of the vegetation, the dominant species, and the canopy openness. Because these other variables were used in the models, canopy openness was superfluous. More detailed canopy cover data that accounts for all levels of the canopy (e.g., Leaf Area Index) may improve the usefulness of this variable.

For similar reasons, the canopy closure classification of the satellite imagery was not particularly useful to the tree models. It appeared in only one tree (the Imagery-Based model for all sites) and it did not provide a very clean split between permafrost probable and improbable. In fact, both nodes were classified as permafrost improbable, but one node continued to be split further with other variables. Because the canopy closure classification is based on the canopy openness data, the two variables are closely related. While the canopy openness variable is closely related to dominant species because of the inherent link to vegetation height, the canopy closure classification is likely closely related to the vegetation classification. Bare ground, tundra, and shrub classes all have open canopies, while the tree classes (broadleaf and needle-leaf) may have either open or closed canopies. This relationship may explain the absence of the canopy closure classification in the tree models when both this classification and the vegetation classification are input into the model.

As discussed in chapter 2, canopy closure and permafrost presence or absence are related (e.g., Brown, 1966; Shur and Jorgenson, 2007; Viereck, 1973). However, an open (or closed) canopy can be an indicator of both permafrost presence and absence depending on what aspect of the relationship is being considered (Brown, 1966; Brown and Péwé, 1973; Geiger *et al.*, 2003; Harding *et al.*, 2002; Heggem *et al.*, 2006; Peddle and Franklin, 1993; Shur and Jorgenson, 2007; Smith and Riseborough, 2002; Viereck, 1973; Yi *et al.*, 2007). This lack of clarity in the relationship between permafrost presence or absence and canopy closure may explain why the canopy closure variables were not useful in the decision tree models.

6.1.2 Vegetation Height

Like the canopy closure classification, vegetation height only appeared in one tree (the Field-Based model). It was used in this tree to extract low shrub and tundra vegetation (under 38 cm tall) from higher shrubs and trees. This variable was replaced by TPI in the tree that included both field and DEM-derived data. Vegetation height is closely related to the dominant species and the class of vegetation at a particular site. Forests have higher vegetation heights than tundra or shrub-dominated sites. In the case of this variable, this close relationship is likely to be one of the reasons for its lack of usefulness. As well, forests are associated with both permafrost presence and absence depending on a number of factors, including species and soil conditions (Beilman and Robinson, 2003; Bliss, 1978; Brown, 1970, 1973; Brown and Péwé, 1973;

Duchesne *et al.*, 2003; Rosenberg *et al.*, 1983; Tryon and Chapin, 1983; Viereck, 1973). Although stunted growth is often associated with permafrost presence, there are many other reasons, unrelated to permafrost presence, why a forest stand may not be particularly tall. In the case of this research, a very open stand may result in a lower value recorded for vegetation height (if the tallest plant at several points was lower, but the dominant species was still a tree species). This unclear division between permafrost presence and absence based on vegetation height is supported by the results of the t-test, in which the differences between vegetation heights at sites with and without permafrost were not significant. The fact that forest sites, including all species, stages of succession, and ultimately heights, and tundra sites are all found at permafrost sites also could explain why vegetation height is not necessarily a useful variable. Furthermore, as discussed in chapter 2, stunted black spruce are often an indicator of permafrost presence (Duchesne *et al.*, 2003; Shur and Jorgenson, 2007; Viereck, 1973); however, black spruce forests were removed from the classification in an earlier split in the tree, prior to the split based on vegetation height (see p. 67, Figure 5.4). As a result, vegetation height only separated tundra vegetation and low shrubs from the taller shrubs and trees. This separation has no clear link to permafrost presence or absence, as demonstrated by both nodes being classified as permafrost improbable.

6.1.3 TPI

TPI also appears to have an unclear relationship to permafrost presence or absence. It appears in seven trees. Most notably, it appears in all three trees created using all study sites and including DEM-derived variables. The insignificant difference between permafrost presence and absence with regard to TPI can be explained by the fact that the distributions are not normal. However, an examination of the histogram comparing TPI distributions for permafrost presence and absence reveals that the two distributions follow similar patterns (Figure 6.1). This indicates that any use of TPI to differentiate between permafrost presence and absence will be based on the data used for that comparison rather than a more universal relationship between TPI (particularly as a proxy for cold air drainage) and permafrost presence or absence. This is also supported by the lack of transferability of the variable, as clearly demonstrated in the DEM-Based trees for Dawson and Faro. In Dawson, sites with TPI values less than 2 were classified as permafrost probable, while in Faro, sites with TPI values less than 2.5 were classified as permafrost improbable.

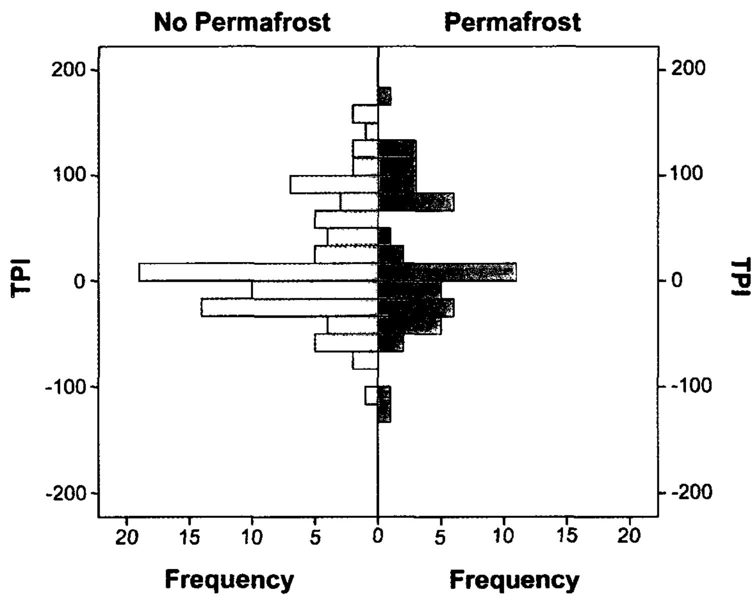


Figure 6.1: Histogram comparing permafrost present and absent TPI data.

6.1.4 Organic Mat Thickness

Organic mat thickness appears about as frequently as TPI appears in the trees. Its relationship between permafrost presence and absence, however, is slightly more clear and predictable. In each tree except the field data only tree for Faro, organic mat thickness is split in the same way: thinner than 10.5 cm is classified as permafrost, between 10.5 cm and 22.5 cm is classified as not permafrost, and thicker than 22.5 cm is classified as permafrost. This is consistent with previous research that found that thicker organic mats associated with black spruce were also associated with permafrost presence (Beilman and Robinson, 2003; Bliss, 1978; Brown, 1966, 1973, 1978; Cowell *et al.*, 1978; Viereck, 1973, 1982). The thinnest organic mats are often associated with tundra vegetation, which in many cases is also associated with permafrost presence. Thin organic mats are also associated with broadleaf vegetation where decomposition is accelerated, moss development is limited, and permafrost is most often absent (Aber and Melillo, 1991; Bliss, 1978; Brown, 1966, 1973; Leverington, 1995; Van Cleve *et al.*, 1983; Van Cleve and Viereck, 1983; Viereck *et al.*, 1983). The unclear division between permafrost presence and absence at sites with thinner organic mats explains why this variable is not used as a primary split. Rather, it is used after the dominant species data has been extracted.

6.1.5 Vegetation Classification

The vegetation classification appeared in three trees: as the primary split for the Imagery-Based model for Keno and the Imagery-Based model for all sites (Figure 5.8 C, p. 78, and 5.7, p. 75, respectively), and as the third split in the Imagery and DEM-Based model for all data (see p. 81, Figure 5.9). The variable is surpassed in usefulness by DEM-derived variables, as is very evident in the case of the Keno trees. The primary split switches to PISR when DEM-derived variables are added to the model, and the accuracy increases nearly 20% from 77.3% to 95.5% (see p. 84, Figure 5.10 C). The inability of the vegetation classification to provide a clear separation between permafrost presence and absence is likely due to the broad classifications. In particular, the "needle-leaf" classification contains species that are associated with both permafrost presence (e.g., black spruce) and absence (e.g., white spruce or lodgepole pine) (Beilman and Robinson, 2003; Bliss, 1978; Brown, 1973; Duchesne *et al.*, 2003; Leverington, 1995; Viereck, 1973). The studies by Morrissey *et al.* (1986) and Duchesne *et al.* (2003) found more detailed vegetation classifications to be useful for permafrost modelling. However, both studies used existing detailed vegetation maps to vaildate the classifications. Detailed vegetation maps are not available for the study area of this project.

6.1.6 Elevation

The use of elevation in the models, like the vegetation classification variable, is consistent with previous research (e.g., Gruber and Hoelzle, 2001; Lewkowicz and Bonnaventure, 2008; Ødegård *et al.*, 1999). Higher elevations tend to be classified as probable permafrost. Lower elevations, however, are more ambiguous, likely due to cold air drainage as well as the vegetation characteristics at that site. As discussed in section 2.5 (p. 17), permafrost that is being protected by vegetation may be preserved at lower elevations where air temperatures may not be conducive to the development of permafrost under current climatic conditions (Shur and Jorgenson, 2007). In the models for Faro, the elevation splits result in opposite predictions (no permafrost at higher elevations). This may be explained by misclassification of permafrost present and absent sites in the field due to very thick active layers or bedrock sites at high elevations.

6.1.7 NDVI

Although NDVI appears as the most useful of the imagery-derived variables in the tree models, its individual contribution to the models is small. In contrast to the vegetation classification variable, in two cases NDVI continues to be included in the models after DEM-derived variables have been added. However, models for individual study areas that include NDVI benefit from an improvement in accuracy over the DEM-Based model only in one location (Johnson's Crossing). In all other cases, the models with NDVI are actually less accurate than the DEM-Based model. Moreover, NDVI is only weakly correlated with the field data, and it does not appear to be related to the dominant species at a site. This indicates that NDVI is not acting as a surrogate, or certainly not as a very good surrogate, for actual vegetation characteristics at a site.

There are significant differences in NDVI values between permafrost and no permafrost sites, and the variable does appear to be useful, particularly when

DEM-derived variables are not involved in the model. However, it does not appear to provide a considerably better classification of permafrost presence and absence than the DEM-derived data alone. The limited contribution of NDVI to the classification is consistent with results of past research in which a vegetation index provided only a limited increase in the quality of the model (Gruber and Hoelzle, 2001; Ødegård *et al.*, 1999). The high correlation found in the European studies between the vegetation index and elevation (Gruber and Hoelzle, 2001; Ødegård *et al.*, 1999) was not found in the Yukon study areas. The coefficient of correlation between NDVI and elevation in this research was significant but low (-0.174) in comparison to -0.56 found in Switzerland by Gruber and Hoelzle (2001) and -0.76 found in Norway by Ødegård *et al.* (1999). This difference is likely a result of the fact that the European studies were focussed on areas with permafrost primarily above treeline.

6.1.8 PISR

The importance of PISR in the tree models is demonstrated through its inclusion in all models that used all sites and in several of the individual study area models. In the models created from all sites, PISR is the primary split except in the model that includes field data. In that model, the importance of PISR is surpassed by the data collected in the field. Solar radiation variables have been useful in numerous studies in the past (e.g., Gruber and Hoelzle, 2001; Leverington and Duguay, 1996; Lewkowicz and Bonnaventure, 2008; Lewkowicz and Ednie, 2004), so its value here is not surprising.

6.1.9 Dominant Species

Of the data collected in the field, the most useful for modelling permafrost distribution was the dominant plant species at a site. There is no imagery-derived variable that can act as a surrogate for the dominant species. All models for which the dominant species was an input variable included it as the primary split. No DEM-derived variable surpassed dominant species in its ability to differentiate between permafrost presence and absence. An interesting point to note, however, is the differences in the actual splits of the dominant species. When dominant, black spruce is always classified as permafrost, but other dominant species classified as permafrost differ among the individual areas. These findings are consistent with past research in which black spruce is the primary species identified with permafrost presence (Beilman and Robinson, 2003; Bliss, 1978; Duchesne *et al.*, 2003; Leverington, 1995; Viereck, 1973). Similarly, where trembling aspen or lodgepole pine were dominant, the site was always classified as permafrost improbable, regardless of study area. All other dominant species (with the exception of black spruce) were classified as permafrost probable in some study areas and permafrost improbable in others.

Additional species classified as permafrost probable in this study may be associated with varying successional stages due to disturbance (Aber and Melillo, 1991; Shur and Jorgenson, 2007; Viereck *et al.*, 1983). Increasing organic mat thickness, decreasing active layer thickness, and increasing moss ground cover are often associated with black spruce forests (Aber and Melillo, 1991; Beilman and Robinson, 2003; Bliss, 1978; Brown, 1966, 1973, 1978; Cowell *et al.*, 1978; Van Cleve *et al.*, 1983; Van Cleve and Viereck, 1983; Viereck, 1973, 1982; Viereck *et al.*, 1983). This can result in a competition for nutrients between the moss and the black spruce trees (Aber and Melillo, 1991; Van Cleve *et al.*, 1983; Van Cleve and Viereck, 1983; Viereck *et al.*, 1983). Therefore, a very late successional stage may be more moss-dominated than black spruce-dominated (Aber and Melillo, 1991; Van Cleve *et al.*, 1983; Viereck *et al.*, 1983). Because many plant species are present throughout all successional stages, mislabelling of the dominant species is likely. In many sites with very thick moss layers, highly scattered black spruce, and scrub birch in most open spaces, the dominant species was often labelled as scrub birch. This species is not necessarily associated with permafrost presence, particularly where dense, tall shrubs dominate; thus, the classification did not clearly represent the characteristics of the vegetation at that site with regard to previously discussed relationships between plant species and permafrost presence or absence.

It seems likely that a single dominant species for any given site is not adequate for determining the vegetation characteristics of the location. White spruce is another species that alone, does not clearly indicate either permafrost presence or absence. But when combined with more detailed site descriptions, it can more accurately predict permafrost presence or absence. For example, white spruce-dominated sites were found in conjuction with well-drained soils, broadleaf trees, very thin organic mats, and an absence of permafrost, as well as with poorly-drained soils, black spruce trees, very thick organic mats, and permafrost. Thus, permafrost distribution can be modelled to some degree using a dominant species classification, particularly in small study areas, but the accuracy of the model is limited due to potential mislabelling of a site or ambiguity related to the permafrost classification and any given species.

For permafrost modelling using field data, a classification that takes into account soil and moisture conditions, all plant species present, and perhaps successional stage (by examining the history of a site) might be useful. Because this study focussed on attempting to use satellite imagery for the vegetation data, such classifications were not used. Instead, a single dominant species label was selected to validate the imagery classification, which is based on the strongest species signature over any given pixel. Thus the dominant species labels were intended to represent, in a more detailed fashion, the classes selected for the remotely sensed vegetation classification.

6.2 Transferability

In general, higher overall accuracies are found using the regression trees for individual study areas than in the trees using all data. This occurs for every area in relation to trees using only field data, for three areas for the trees using field data and DEM-derived data, the trees using imagery-derived and DEM-derived data, and the trees using only imagery-derived data, and in four of the five areas for the DEM-Based trees. As well, for each tree model, the average overall accuracy for the individual area was higher in all cases than the overall accuracy for the tree created using all data (Table 6.1). This indicates that the relationship between permafrost and the variables in a given area is not completely transferrable to other areas. When the all-data trees were created, most had lower accuracies than the respective trees for individual areas.

Table 6.1: Mean overall accuracy for trees created for individual areas compared to overall accuracy for each tree created using all data.

	Mean Overall Accuracy (Individual Areas)	Overall Accuracy (All Data)
DEM-Based	91.4%	88.2%
Field-Based	90.2%	84.3%
Field and DEM-Based	91.5%	89.2%
Imagery-Based	76.6%	74.5%
Imagery and DEM-Based	90.7%	90.2%

This limited transferability is further emphasized by the variables used for individual study areas compared to the trees with all data. For the trees that used data collected in the field, the dominant species was used for the primary (and in most cases only) split. The likely reason for the differences in overall accuracy (and the lower overall accuracy in the all-data model) is the fact that this variable is being split differently in each area. Therefore, when combined for all study areas, it is less effective because the group of species classified as being associated with permafrost differs from one area to another. Black spruce always results in a node classified as permafrost, but the other species classified as permafrost are not consistent.

For the trees that use imagery-derived variables only (see p. 78, Figure 5.8), the primary splits are created using either NDVI or the vegetation classification. For two sites (Dawson and Faro), NDVI was used with different breakpoints, but the classifications were similar (higher NDVI values indicated no permafrost while lower NDVI values indicated permafrost presence). This is consistent in Keno, where the primary split is created using the vegetation class, but the second split uses NDVI in a similar manner to the trees for Dawson and Faro (higher NDVI values representing no permafrost). This similar use of the imagery-derived variable for Keno, Dawson, and Faro also indicates that the different image dates used did not affect the results. Although this use of NDVI is consistent throughout the individual areas, when the variable is included in the tree with all data, it no longer follows a simple relationship. The tree with all data (see p. 75, Figure 5.7) uses NDVI for a number of splits; in all cases, the highest values are classified as permafrost improbable, but the lower values are not consistently classified as permafrost probable. On one branch of the tree, the highest and the lowest NDVI values indicate no permafrost, but the middle values are classified as permafrost. This inconsistency is also demonstrated in the trees that contain imagery-derived data and DEM-derived data (see p. 84, Figure 5.10). In the two trees that include NDVI (Faro and Johnson's Crossing), the use of NDVI differs. In Faro, as expected based on the majority of cases, lower NDVI values indicate permafrost presence while higher values indicate no permafrost (although a further split using TPI extracts a permafrost class). In Johnson's Crossing, however, the relationship is the opposite: lower NDVI values represent no permafrost while higher values indicate permafrost presence. This also occurs in the tree with all the data (see p. 81, Figure 5.9). In most cases higher NDVI values represent no permafrost, except in one split where the higher values indicate permafrost presence.

Interestingly, the transferability problem is not unique to the variables that represent vegetation characteristics. The DEM-Based trees (with no vegetation data) also have higher overall accuracies for all the individual areas except for Sa Dena Hes where no split was created. This could be caused by a number of factors. First, no variable comes out as being most important for the individual areas. Elevation provides the primary split for two areas (with different break points), while PISR provides the primary split for two other areas (this time with similar break points). The tree with all data has PISR as the primary split, but the break point is considerably higher than either of the two individual models that use it. As well, there may be some over-fitting with the trees for individual areas that does not appear in the tree with all data. This would result in inflated overall accuracies for the individual study areas and a lower, perhaps more realistic, overall accuracy for the tree with all the data. This over-fitting may also be a problem for the trees created using vegetation variables; however, in many cases the trees contain only one split, so over-fitting is less likely.

6.3 Maps

The identification of the best maps for each study area, selected based on the most accurate CART model, produced interesting results. The majority of selected maps were created from tree models using only data from individual study areas. Only the Imagery and DEM-Based tree using all field sites was the best map in Sa Dena Hes. This emphasis on the trees created from data from individual study areas is further support for the lack of transferability of the relationship between vegetation data and permafrost presence or absence across large areas.

The percentage of permafrost coverage for each map was not always representative of the permafrost zone as defined by Heginbottom *et al.* (1995). For Faro, Keno, and Johnson's Crossing, the permafrost coverage in the majority of the maps fell within the range defined for the permafrost zone in which each area was located. For Dawson and Sa Dena Hes, however, only the map produced from the DEM-Based model had a percentage of permafrost coverage within the range defined for the permafrost zone. The significantly smaller scale of the map produced by Heginbottom *et al.* (1995) likely explains these differences in permafrost coverage.

6.4 Errors

Errors in the data could have affected the resulting models. Manual inputting of data is certainly a potential source of error, but all data were checked to reduce, and ideally eliminate, any errors. Two additional potential sources of error include imprecision of the coordinates of study sites recorded with the GPS and inaccuracy in the classification of the satellite imagery.

6.4.1 Site Location Errors

Errors associated with the coordinates recorded for the field sites and how these relate to the pixels in the imagery are one important factor to consider. The accuracy of the GPS ranged from about 5 m up to 20 m (and 50 m at a very small number of sites). This inaccuracy means that an indicated site could be in a different location and, more importantly, could actually belong in a different pixel from the one in which it was placed for this research. Using the range of 5 to 20 m of error, the actual location of the site depends on where the recorded point is within its cell and how inaccurate the coordinates truly are; for 20 m of error, the actual point could be anywhere within a circle with a radius of 20 m around the recorded point. A given coordinate could, therefore, actually be located in any of nine adjacent pixels. The outer point of the transect could also be located only within those nine pixels. If the transects had been any longer (e.g., 30 m instead of 20 m), they could have extended into any of 21 adjacent pixels. Fortunately, in the error assessments for the vegetation and canopy closure classification of the imagery, the potential error caused by misplacement of points is minimized because in the majority of cases the classification at a point was the same for surrounding pixels. As well, sites were selected based on surrounding homogeneity; sites were not examined at the edge of a stand. Around any given site, a buffer of at least 30 m of similar vegetation characteristics was left to reduce classification errors associated with the inaccuracies of the recorded coordinates.

6.4.2 Image Classification Errors

Misclassified pixels in the canopy closure and vegetation classifications of the satellite imagery may also have had an effect on how valuable each variable is in the tree models. For the vegetation classification and the canopy closure, approximately 30% of the pixels are misclassified. These errors, once carried over into the tree models, can limit the ability of the variables to predict permafrost distribution. This also indicates that vegetation or canopy closure classifications created from Landsat imagery may not be reliable for permafrost modelling. As well, the classes that can be created for the vegetation classification are not useful for permafrost modelling, as discussed in section 6.1 (p. 97). Duchesne *et al.* (2003) separated the different needle-leaf species because they are indicators of opposing ground conditions. To create this more detailed classification, the researchers used a combination of a pre-existing vegetation classification map and Landsat imagery. This separation of species was not possible for this research because vegetation data at a large enough scale were not available for the study areas.

6.5 Best Models

The "best" model differs depending on the aspects being considered, such as transferability, testing accuracy, training accuracy, or included variables.

The Field and DEM-Based tree had an increase in accuracy over the DEM-Based tree for both the training and the testing data. As well, this model included the most useful variable overall, dominant species. Although the Imagery and DEM-Based tree had a higher training accuracy, the considerably higher testing accuracy of the Field and DEM-Based tree indicates that it is, overall, the best model. The primary disadvantage of any model that included vegetation data collected in the field is that permafrost distribution could not be accurately modelled over the study areas. Remotely sensed vegetation data could not act as a proxy for the field data, so a continuous coverage of each variable could not be created.

An important point to note is that the imagery-derived data alone produced a tree with the lowest overall accuracies, but, when these variables were combined with the DEM-derived data, they produced a tree with the highest overall accuracy. There are at least two possible explanations for this result. First, although the imagery-derived variables on their own were the least capable of accurately predicting permafrost distribution, it is possible that they provide enough different information from the DEM-derived data that they increase the predictive capability of the DEM-Based model. The weak correlation between NDVI and elevation and the insignificant correlations between NDVI and PISR and TPI indicate that this is a possible explanation. In addition, stronger correlations between NDVI and canopy openness and vegetation height indicate that NDVI might be contributing some vegetation data that cannot be represented as clearly by the DEM-derived data. Second, the Imagery and DEM-Based tree may include some over-fitting of the data. Over-fitting would also explain why the Imagery and DEM-Based tree had the highest training accuracy but the second lowest testing accuracy. The Imagery-Based tree was never selected as the best model based on the maps, while the Imagery and DEM-Based tree was selected as the best model only for Sa Dena Hes. Therefore, although it had the highest accuracy with the training data, its low testing accuracy and the fact that it only resulted in one "best" map indicate that it does not produce a true representation of permafrost distribution across the southern Yukon.

Based on the accuracies of the models, the DEM-Based trees and Imagery and DEM-Based trees for individual study areas were found to produce the best maps. The accuracies for the DEM-Based tree and the Imagery and DEM-Based tree were identical for each study area (because the models were identical) except for Johnson's Crossing, for which the Imagery and DEM-Based tree had a higher accuracy. Faro, however, had a lower accuracy with the Imagery and DEM-Based tree model.

These results indicate that, while vegetation variables alone, both collected in the field and collected from satellite imagery, can be used to model presence or absence of permafrost, DEM-derived variables alone can produce more accurate models. The addition of remotely sensed vegetation data to the DEM-derived data can produce a model that is more accurate for the sites used for creating the model; however, this method is less accurate when different sites are classified using the model. Data from an individual area can produce more accurate models for that area. However, inclusion of remotely sensed vegetation data in such models often does not improve the result.

6.6 Practical Use for Models

Permafrost is generally avoided in the construction of buildings or roads. Therefore, accurate modelling of permafrost absence would be most useful for practical applications. While there is variability in the user's accuracies for the permafrost probable classification, most of the models show higher than 80% user's accuracies for the permafrost improbable classification (Appendix E).

The DEM-Based and Imagery and DEM-Based models for individual study areas had user's accuracies for the permafrost improbable class ranging from 86.7% to 100%. This is consistent with the models containing all data, in which the DEM-Based model and the Imagery and DEM-Based model had the highest user's accuracies for the permafrost improbable class in the training data (89.6% and 92.3%, respectively). The DEM-Based model had the highest testing accuracy of the permafrost improbable class (80%).

These models could be included as a variable for modelling routes for road construction. When there is even a low margin of error, ground-truthing is necessary to confirm the absence of permafrost. However, with greater than 80% accuracies, these models provide a good basis for selecting potential routes.

7 CONCLUSION

This research examined the use of variables representing vegetation in CART models of permafrost distribution. Three groups of variables, ten variables in total, were tested in different combinations to determine the value of vegetation variables as predictors. Although other studies have considered these variables, only four have been tested in permafrost models. A number of previous studies have tested elevation and solar radiation variables in permafrost models (e.g. Gruber and Hoelzle, 2001; Leverington and Duguay, 1996; Lewkowicz and Bonnaventure, 2008; Ødegård *et al.*, 1999). Several studies have also tested the viability of including NDVI and land cover classifications in permafrost models (e.g. Duchesne *et al.*, 2003; Gruber and Hoelzle, 2001; Morrissey *et al.*, 1986; Ødegård *et al.*, 1999). While the relationship between permafrost distribution and vegetation data collected in the field has been examined, no study has directly included these vegetation variables in permafrost models (e.g. Brown, 1966; Brown and Péwé, 1973; Van Cleve *et al.*, 1983; Viereck, 1973). In this study, individual variables were also examined to determine the most useful variable for representing vegetation when modelling permafrost. Finally, models for each individual study area were created to compare the transferability of each variable with regard to permafrost distribution.

The results showed that models containing only vegetation variables (either field-derived or imagery-derived) can differentiate between permafrost probable and improbable sites, but the accuracies are not as high as for models containing only DEM-derived variables (88.2% training accuracy, 73.5% testing). Vegetation variables produced accuracies of 84.3% (70.6% testing) for field-derived variables and 74.5% (55.9% testing) for imagery-derived variables. The addition of each group of vegetation variables to the DEM-derived variables increased the overall training accuracies, but only minimally and the testing accuracies were inconsistent with the training accuracy increases. The combination of DEM-derived and field-derived vegetation variables showed an increase in accuracy over the DEM-Based tree model (89.2% training, 76.5% testing). This model, however, requires extensive fieldwork to collect the data and cannot be modelled remotely. Therefore, it can be used either as a point indicator of permafrost presence or absence at sites with known vegetation characteristics or as a generalization of the relationship between permafrost and vegetation. The addition of imagery-derived variables to the DEM-derived variables produced an increase in training accuracy, but a large decrease in testing accuracy compared to the DEM-Based model (90.2% training, 64.7% testing). The lower testing accuracy of this model suggests that there was some over-fitting with this model and that the DEM-Based model alone is less complicated and actually more accurate than when imagery-derived vegetation variables are added. The increase in training accuracy is minimal and there is an increased risk of over-fitting in the models; thus, it is likely that the additional cost and time required for incorporating the vegetation variables into the models may exceed the value of the variables to the model.

The most useful vegetation variable was found to be the dominant species. This variable provided the clearest and most accurate differentiation of permafrost probable and permafrost improbable sites. However, this variable had several limitations. First, the only species that could be extracted as clearly relating to permafrost presence was black spruce, while the only species that always related to an absence of permafrost were lodgepole pine and trembling aspen. All other species were dominant both at sites with and without permafrost; therefore, a clear classification for all other species was not possible. Additionally, the dominant species could not be modelled using satellite imagery at the scale of the permafrost models for this study. Duchesne *et al.* (2003) did create an Enhanced Vegetation Classification in the Mackenzie River valley for which detailed vegetation maps are available; however, such maps are not available for the Yukon. Higher spatial resolution imagery may help with these classifications, but such imagery is very expensive. Finally, even if imagery can be classified by species, it cannot be used over large areas because of the variability of the relationship between permafrost and species across different regions.

In most cases, accuracies for the tree models created using all data were high. However, examination of the models created for individual study areas showed that the use of the variables differed considerably between study areas. The inconsistency with variable use across areas, and the higher accuracies for models created from data for individual study areas as compared to models created from data from all areas indicates a lack of transferability of the relationship between vegetation and permafrost. The models for Faro and Dawson, study areas within the same climatological region, were dissimilar in their structure and use of variables. The models for Faro and Keno, within the same ecoregion, also differed in their complexity and use of variables. These variations show that the relationships between vegetation and permafrost presence or absence differ over distances that are smaller than climatological regions or ecoregions.

The results of this research indicate that models of permafrost distribution across large areas are not likely to be enriched sufficiently by the inclusion of vegetation variables. On the other hand, permafrost models for small study areas may benefit from the inclusion of vegetation data if a distinct relationship can be found at that location. These results support the finding of Leverington and Duguay (1997) that their permafrost model near Mayo, Yukon using land cover and equivalent latitude was not transferrable to a second study area less than 50 km away. Further examination of the variation in the relationship between vegetation and permafrost at different scales would be required to determine at which scale vegetation is best suited for permafrost models. Additionally, more detailed classifications that account for all levels of vegetation cover at each site (ground cover, shrubs, trees, and canopy closure) may help account for variations found with the dominant species variable.

The models produced for this project were limited to a binary output (permafrost probable or improbable). Because of the method used, more detailed output would likely have reduced the reliability of the models in this research. However, the CART method for permafrost modelling has not been used previously, thus further testing of this method may help refine the results. Further examination of additional methods (e.g., logistic regression) for including vegetation in permafrost distribution models could also be tested for the quantitative variables to produce a more detailed output model.

Finally, predicting where permafrost is absent in the discontinuous permafrost zone is usually the key task for infrastructure development. The majority of the CART models produced in this research can predict permafrost improbable areas with user's accuracies greater than 80%. This high accuracy and the fact that the best models for classifying areas with no permafrost include vegetation data derived from remotely sensed imagery (the DEM-Based models and the Imagery and DEM-Based models) indicate that these models could provide a good basis for selecting sites and routes for development.

8 REFERENCES

Aber, J.D., & Melillo, J.M. (1991). *Terrestrial Ecosystems*. Philadelphia: Saunders College Publishing.

Beilman, D.W., & Robinson, S.D. (2003). Peatland permafrost thaw and landform type along a climatic gradient. In M. Phillips, S.M. Springman, & L.U. Arenson (Eds.), *Proceedings of the 8th International Conference on Permafrost, Zurich, Switzerland* (pp. 61-65). Lisse, The Netherlands: Swets and Zeitlinger.

Bliss, L.C. (1978). Vegetation and revegetation within permafrost terrain. In *Third International Conference on Permafrost: Proceedings, vol. 2* (pp. 31-50). Ottawa: National Research Council of Canada.

Breiman, L., Friedman, J.H., Olshen, R.A., & Stone, C.J. (1998). *Classification and Regression Trees*. New York: CRC Press.

Brenning, A., Gruber, S., & Hoelzle, M. (2005). Sampling and statistical analyses of BTS measurements. *Permafrost and Periglacial Processes 16*, 383-393.

Brockman, C.F. (2001). *Trees of North America: A Guide to Field Identification*. NewYork: St. Martin's Press.

Brown, N., Jennings, S., Wheeler, P., & Nabe-Nielson, J. (2000). An improved method for the rapid assessment of forest understory light environments. *Journal of Applied Ecology 37*, 1044-1053.

Brown, R.J.E. (1966). *Influence of vegetation on permafrost*. Research Paper No. 298, Division of Building Research, National Research Council of Canada, Ottawa.

Brown, R.J.E. (1970). *Permafrost in Canada*. Toronto: University of Toronto Press.

Brown, R.J.E. (1973). Influence of climatic and terrain factors on ground temperatures at three locations in the permafrost region of Canada. In *Permafrost: Second International Conference* (pp. 27-34). Washington, DC: National Academy of Sciences.

Brown, R.J.E. (1978). Influence of climate and terrain on ground temperatures in the continuous permafrost zone of northern Manitoba and Keewatin district, Canada. In *Third International Conference on Permafrost: Proceedings, vol. 1* (pp. 15-21). Ottawa: National Research Council of Canada.

Brown, R.J.E., & Péwé, T.L. (1973). Distribution of permafrost in North America and its relationship to the environment: a review, 1963-1973. In *Permafrost: Second International Conference* (pp. 71-100). Washington, DC: National Academy of Sciences.

Carey, S.K., & Woo, M.-K. (2001). Spatial variability of hillslope water balance, Wolf Creek basin, subarctic Yukon. *Hydrological Processes 15*, 3113-3132.

Clark, M.J., Gurnell, A.M., Milton, E.J., Seppälä, M., & Kyöstilä, M. (1985). Remotely-sensed vegetation classification as a snow depth indicator for hydrological analysis in sub-arctic Finland. *Fennia 163* (2), 195-216.

Cowell, D.W., Jeglum, J.K., & Merriman, J.C. (1978). Preservation of seasonal frost in peatlands, Kinoje Lakes, southern Hudson Bay Lowland. In *Third International Conference on Permafrost: Proceedings, vol. 1* (pp. 453-459). Ottawa: National Research Council of Canada.

Dingman, S.L., & Koutz, F.R. (1974). Relations among vegetation, permafrost, and potential insolation in central Alaska. *Arctic and Alpine Research 6* (1), 37-47.

Duchesne, C., Wright, J.F., Côté, M.M., & Nixon, F.M. (2003). An enhanced vegetation classification for permafrost modelling, NWT, Canada. In M. Phillips, S.M. Springman, & L.U. Arenson (Eds.), *Proceedings of the 8th International Conference on Permafrost, Zurich, Switzerland* (pp. 217-221). Lisse, The Netherlands: Swets and Zeitlinger.

Dutch, S. (2005). *UTM Conversion Tool* [*.xls File]. University of Wisconsin, Green Bay. Retrieved from: www.uwgb.edu/dutchs/UsefulData/UTMConversions1.xls [December 1, 2008].

Environment Canada. (2009a). *Narrative Decsiption of Terrestrial Ecozones and Ecoregions of Canada.* [Website] http://www.ec.gc.ca/soer-ree/English/framework/nardesc/borcor_e.cfm [September 15, 2009].

Environment Canada. (2009b). National Climate Data and Information Archive. [Website] http://www.climate.weatheroffice.ec.gc.ca/Welcome_e.html [September 29, 2009].

Etzelmüller, B., Heggem, E.S.F., Sharkhuu, N., Frauenfelder, R., Kääb, A., & Goulden, C. (2006). Mountain permafrost distribution modeling using a mutli-criteria approach in the Hövsgöl area, northern Mongolia. *Permafrost and Periglacial Processes 17*, 91-104.

Geiger, R., Aron, R.H., & Todhunter, P. (2003). *The climate near the ground.* (6th ed.). Lanham: Rowman & Littlefield Publishers.

Geomatics Yukon. (2006). *30 Meter Yukon Digital Elevation Model* [data file]. Whitehorse, Yukon. Available download: ftp://ftp.geomaticsyukon.ca/DEMs/30m/ [compiled in 2007].

Government of Yukon, Wildland Fire Management. (2008). *Fire History* [Data File]. Whitehorse, Yukon: Government of Yukon. Available download: http://www.geomaticsyukon.ca/data_download.html [September 9, 2009].

Gray, J.T., & Brown, R.J.E. (1982). The influence of terrain factors on the distribution of permafrost bodies in the Chic-Choc Mountains, Gaspésie, Québec. In H.M. French (Ed.), *Proceedings: Fourth Canadian Permafrost Conference* (pp. 23-35). Ottawa, Canada: National Research Council of Canada.

Gruber, S., & Hoelzle, M. (2001). Statistical modelling of mountain permafrost distribution: local calibration and incorporation of remotely sensed data. *Permafrost and Periglacial Processes 12*, 67-77.

Harding, R., Kuhry, P., Christensen, T.R., Sykes, M.T., Dankers, R., & van der Linden, S. (2002). Climate feedbacks at the tundra-taiga interface. *Ambio Special Report 12*, 47-55.

Heggem, E.S.F., Etzelmüller, B., Anarmaa, S., Sharkhuu, N., Goulden, C.E., & Nanadinsetseg, B. (2006). Spatial distribution of ground surface temperatures and active layer depths in the Hövsgöl area, northern Mongolia. *Permafrost and Periglacial Processes 17*, 357-369.

Heginbottom, J.A., Dubreuil, M.A., & Harker, P.T. (1995). *Canada Permafrost.* (1:7,500,000 scale). The National Atlas of Canada, 5th Edition, sheet MCR 4177. Ottawa: National Resources Canada.

Jenness, J. (2006). *Topographic Position Index extension for ArcView 3., v. 1.3a* (tpi_jen.avx). Jenness Enterprises. Available at: http://www.jennessent.com/arcview/tpi.htm [July 29, 2009].

Jensen, J.R. (2005). *Introductory Digital Image Processing.* Toronto: Pearson Education Inc.

Kneisel, C. (2003). Sporadic and discontinuous mountain permafrost occurrence in the Upper Engadine, eastern Swiss Alps. In M. Phillips, S.M. Springman, & L.U. Arenson (Eds.), *Proceedings of the 8th International Conference on Permafrost, Zurich, Switzerland* (pp. 561-566). Lisse, The Netherlands: Swets and Zeitlinger.

Kokelj, S.V., & Burn, C.R. (2003). "Drunken forest" and near-surface ground ice in Mackenzie Delta, Northwest Territories, Canada. In M. Phillips, S.M. Springman, & L.U. Arenson (Eds.), *Proceedings of the 8th International Conference on Permafrost, Zurich, Switzerland* (pp. 567-571). Lisse, The Netherlands: Swets and Zeitlinger.

Lee, R. (1978). *Forest Microclimatology.* New York: Columbia University Press.

Leverington, D. (1995). A field survey of late-summer depths to frozen ground at two study areas near Mayo, Yukon Territory, Canada. *Permafrost and Periglacial Processes 6*, 373-379.

Leverington, D.W., & Duguay, C.R. (1996). Evaluation of three supervised classifiers in mapping "depth to later-summer frozen ground," central Yukon Territory. *Canadian Journal of Remote Sensing 22* (2), 163-174.

Leverington, D.W., & Duguay, C.R. (1997). A neural network method to determine the presence or absence of permafrost near Mayo, Yukon Territory, Canada. *Permafrost and Periglacial Processes 8*, 205-215.

Lewkowicz, A.G., & Bonnaventure, P.P. (2008). Interchangeability of mountain permafrost probability models, northwest Canada. *Permafrost and Periglacial Processes 19*, 49-62.

Lewkowicz, A.G., & Ednie, M. (2004). Probability mapping of mountain permfrost using the BTS method, Wolf Creek, Yukon Territory, Canada. *Permafrost and Periglacial Processes 15*, 67-80.

Lillesand, T.M., Kiefer, R.W., & Chipman, J.W. (2008). *Remote Sensing and Image Interpretation.* (6th ed.). Hoboken, New Jersey: John Wiley & Sons.

MacKinnon, A., Pojar, J., & Coupé, R. (Eds.). (1999). *Plants of Northern British Columbia.* (2nd ed.). Vancouver: B.C. Ministry of Forests and Lone Pine Publishing.

Meyer, P., Itten, K.I., Kellenberger, T., Sandmeier, S., & Sandmeier, R. (1993). Radiometric corrections of topographically induced effects on Landsat TM data in an alpine environment. *Journal of Photogrammetry and Remote Sensing 48* (4), 17-28.

Morrissey, L.A., Strong, L.L., & Card, D.H. (1986). Mapping permafrost in the Boreal Forest with Thematic Mapper satellite data. *Photogrammetric Engineering and Remote Sensing 52* (9), 1513-1520.

Ødegård, R.S., Isaksen, K., Mastervik, M., Billdal, L., Engler, M., & Sollid, J.L. (1999). Comparison of BTS and Landsat TM data from Jotunheimen, southern Norway. *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography 57*, 226-233.

Oke, T.R. (1987). *Boundary Layer Climates*. (2nd ed.). New York: Methuen.

Peddle, D.R., & Franklin, S.E. (1993). Classification of permafrost active layer depth from remotely sensed and topographic evidence. *Remote Sensing of Environment 44*, 67-80.

Pomeroy, J.W., Bewley, D.S., Essery, R.L.H., Hedstom, N.R., Link, T., Granger, R.J., Sicart, J.E., Ellis, C.R., & Janowicz, J.R. (2006). Shrub tundra snowmelt. *Hydrological Processes 20*, 923-941.

Pomeroy, J.W., & Goodison, B.E. (1997). Winter and Snow. In W.G. Bailey, T.R. Oke, & W.R. Rouse (Eds.), *The surface climates of Canada* (pp. 68-100). Montreal & Kingston: McGill-Queen's University Press.

Ripley, B.D. (1996). *Pattern Recognition and Neural Networks*. Cambridge, UK and New York: Cambridge University Press.

Rosenberg, N.J., Blad, B.L., & Verma, S.B. (1983). *Microclimate: The biological environment*. (2nd ed.). Toronto: John Wiley & Sons.

Rouse, J.W. Jr., Hass, R.W., Schell, J.A., Deering, D.W., & Harlan, J.C. (1974). Monitoring the vernal advancement and retrogradation (Greenwave effect) of natural vegetation. NASA/GSFCT Type III Final Report, Greenbolt, MD, USA.

Rouse, W.R. (1982). Microclimate of low arctic tundra and forest at Churchill, Manitoba. In H.M. French (Ed.), *Proceedings: Fourth Canadian Permafrost Conference* (pp. 68-80). Ottawa, Canada: National Research Council of Canada.

Shur, Y.L., & Jorgenson, M.T. (2007). Patterns of permafrost formation and degradation in relation to climate and ecosystems. *Permafrost and Periglacial Processes 18*, 7-19.

Smith, M.W., & Riseborough, D.W. (2002). Climate and the limits of permafrost: a zonal analysis. *Permafrost and Periglacial Processes 13*, 1-15.

SPSS Inc. (2009a). *SPSS Classification TreesTM 17.0*. SPSS Inc. Retrieved from http://support.spss.com/ProductsExt/SPSS/Documentation/SPSSforWindows/SPSSDecsionTrees17.0.pdf [August 18, 2009].

SPSS Inc. (2009b). *SPSS Statistics Base 17.0 User's Guide.* SPSS Inc. Retrieved from http://support.spss.com/ProductsExt/SPSS/Documentation/SPSSforWindows/ SPSS Statistics Base User's Guide 17.0 [August 1, 2009].

StatSoft, Inc. (2007). *Electronic Statistics Textbook.* Tulsa, OK: StatSoft. Retrieved from http://www.statsoft.com/textbook/stathome.html [August 16, 2009].

Steyn, D.G., Schmid, H-P., Walmsley, J.L., & Wilson, J.D. (1997). Spatial variability in surface climates. In W.G. Bailey, T.R. Oke, & W.R. Rouse (Eds.), *The surface climates of Canada* (pp. 44-67). Montreal & Kingston: McGill-Queen;s University Press.

Teillet, P.M., Guindon, B., & Goodenough, D.G. (1982). On the slope-aspect correction of multispectral scanner data. *Canadian Journal of Remote Sensing 8* (2), 84-106.

Tryon, P.R., & Chapin, F.S. III. (1983). Temperature control over root growth and root biomass in taiga forest trees. *Canadian Journal of Forest Research 13*, 827-933.

US Geological Survey. (2009a). *Landsat TM5 and Landsat ETM+7 Imagery* [Data files]. USGS Available download: http://glovis.usgs.gov and http://earthexplorer.usgs.gov [December 2008 to February 2009].

US Geological Survey. (2009b). *SLC-off Products: Background.* Retrieved from http://landsat.usgs.gov/products_slcoffbackground.php [July 29, 2009].

Van Cleve, K., Dyrness, C.T., Viereck, L.A., Fox, J., Chapin, F.S. III, & Oechel, W. (1983). Taiga ecosystems in interior Alaska. *BioScience 33*, 39-44.

Van Cleve, K., & Viereck, L.A. (1983). A comparison of successional sequences following fire on permafrost-dominated and permafrost-free sites in interior Alaska. In *Proceedings of Fourth International Conference on Permafrost* (pp. 1292-1297). Washington, DC: National Academy Press.

Viereck, L.A. (1973). Ecological effects of river flooding and forest fires on permafrost in the taiga of Alaska. In *Permafrost: Second International Conference* (pp. 60-67). Washington, DC: National Academy of Sciences.

Viereck, L.A. (1982). Effects of fire and firelines on active layer thickness and soil temperatures in interior Alaska. In H.M. French (Ed.), *Proceedings: Fourth Canadian Permafrost Conference* (pp. 123-135). Ottawa, Canada: National Research Council of Canada.

Viereck, L.A., Dyrness, C.T., Van Cleve, K., & Foote, M.J. (1983). Vegetation, soils, and forest productivity in selected forest types in interior Alaska. *Canadian Journal of Forest Research 13*, 703-720.

Wahl, H.E., Fraser, D.B., Harvey, R.C., & Maxwell, J.B. (1987). *Climate of Yukon.* Ottawa, Canada: Canadian Government Publishing Centre.

Walker, J. (1995). *Solar System Live.* Website. http://www.fourmilab.ch/cgi-bin/Solar/action?sys=-Sf [February 2009].

Weiss, A. (2001). *Topographic Position and Landforms Analysis.* Poster presentation, ESRI User Conference, San Diego, CA.

Woo, M.-K., & Marsh, P. (2005). Snow, frozen soils and permafrost hydrology in Canada, 1999-2002. *Hydrological Processes 19*, 215-229.

Yi, S., Woo, M.-K., & Arain, M.A. (2007). Impacts of peat and vegetation on permafrost degradation under climate warming. *Geophysical Research Letters 34*, L16504.

Zhang, T., Osterkamp, T.E., & Stamnes, K. (1997). Effects of climate on the active layer and permafrost on the North Slope of Alaska, U.S.A. *Permafrost and Periglacial Processes 8*, 45-67.

APPENDIX A

Number of dots counted using canopy scope and corresponding percentage of canopy openness. Calculated using Equation 4.1.

Dots	Percent	Dots	Percent	Dots	Percent
0	0.0	33	34.4	67	49.0
0	2.2	34	35.0	68	49.4
1	5.9	35	35.5	69	49.8
2	8.6	36	36.0	70	50.1
3	10.5	37	36.5	71	50.5
4	12.1	38	37.0	72	50.8
5	13.5	39	37.4	73	51.2
6	14.8	40	37.9	74	51.5
7	16.0	41	38.4	75	51.9
8	17.1	42	38.8	76	52.2
9	18.1	43	39.3	77	52.5
10	19.1	44	39.8	78	52.9
11	20.0	45	40.2	79	53.2
12	20.8	46	40.7	80	53.5
13	21.7	47	41.1	81	53.9
14	22.5	48	41.5	82	54.2
15	23.3	49	42.0	83	54.5
16	24.1	50	42.4	84	54.9
17	24.8	51	42.8	85	55.2
18	25.5	52	43.2	86	55.5
19	26.2	53	43.6	87	55.8
20	26.9	54	44.1	88	56.1
21	27.5	55	44.5	89	56.5
22	28.2	56	44.9	90	56.8
23	28.8	57	45.3	91	57.1
24	29.4	58	45.6	92	57.4
25	30.0	59	46.0	93	57.7
26	30.6	60	46.4	94	58.0
27	31.2	61	46.8	95	58.3
28	31.8	62	47.2	96	58.6
29	32.3	63	47.6	97	58.9
30	32.9	64	47.9	98	59.2
31	33.4	65	48.3	99	59.5
32	33.9	66	48.7	100	59.8

APPENDIX B

All variables for each site used for analysis.

ID	Area	Latitude (° N)	Longitude (° W)	Elevation (m)	PISR (MJ/m²)	TPI	Organic Mat Thickness (cm)	Vegetation Height (cm)	Canopy Openness (%)	Dominant Species	NDVI	Vegetation Class	Canopy Cover Class	Permafrost
1	JC	60.5835	133.0809	1186	1456.2	-30	20.2	130.0	100.0	*Salix* spp.	0.596	Shrub	Open	Absent
2	JC	60.5835	133.0833	1188	1456.6	-31	12.2	145.0	100.0	*Betula glandulosa*	0.652	Shrub	Open	Absent
3	JC	60.5821	133.0781	1197	1433.7	-22	30.3	80.0	100.0	*Salix* spp.	0.508	Shrub	Open	Absent
4	JC	60.5649	133.0953	1221	1472.4	-30	11.4	1211.2	13.3	*Abies lasiocarpa*	0.614	Needle-Leaf	Closed	Absent
5	JC	60.5671	133.0983	1335	1517.3	63	7.8	780.7	42.5	*Abies lasiocarpa*	0.657	Needle-Leaf	Closed	Absent
6	JC	60.5715	133.1183	1500	1262.9	96	5.7	5.0	100.0	Lichen	0.418	Tundra	Open	Present
7	JC	60.5721	133.1190	1484	1369.5	77	6.8	14.0	100.0	Lichen	0.404	Tundra	Open	Present
8	JC	60.5161	133.2154	897	1378.9	1	12.1	1211.2	18.6	*Pinus contorta* var. *latifolia*	0.425	Broad-Leaf	Open	Absent
9	JC	60.559	133.1158	1152	1472.3	-67	10.0	167.0	100.0	*Salix* spp.	0.632	Shrub	Open	Absent
10	JC	60.5606	133.1168	1169	1478.8	-67	9.8	1708.4	36.5	*Abies lasiocarpa*	0.604	Needle-Leaf	Closed	Absent
11	JC	60.5627	133.1181	1218	1512.0	-59	11.7	1264.0	17.3	*Abies lasiocarpa*	0.704	Needle-Leaf	Closed	Present
12	JC	60.6327	133.0240	1089	1436.7	10	4.1	780.1	42.3	*Pinus contorta* var. *latifolia*	0.662	Needle-Leaf	Closed	Absent
13	JC	60.5525	133.1443	1141	1470.7	-8	13.3	593.4	48.0	*Abies lasiocarpa*	0.313	Bare	Open	Absent
14	JC	60.6236	133.0342	1123	1450.8	4	11.2	664.3	49.0	*Abies lasiocarpa*	0.721	Needle-Leaf	Closed	Present
15	JC	60.6073	133.0539	1154	1422.8	-20	12.0	165.0	100.0	*Betula glandulosa*	0.366	Bare	Open	Absent
16	JC	60.6078	133.0567	1138	1457.4	-38	34.3	53.0	100.0	*Betula glandulosa*	0.565	Shrub	Open	Absent
17	JC	60.5815	133.0764	1208	1430.1	-14	11.9	838.8	100.0	*Abies lasiocarpa*	0.589	Needle-Leaf	Closed	Absent
18	JC	60.5838	133.0842	1192	1454.5	-28	13.6	780.1	49.8	*Picea glauca*	0.604	Shrub	Open	Absent
19	JC	60.5686	133.1023	1376	1497.1	74	5.8	20.0	100.0	*Salix* spp.	0.428	Tundra	Open	Absent
20	JC	60.5516	133.1438	1135	1460.7	-9	10.9	907.5	21.4	*Abies lasiocarpa*	0.369	Broad-Leaf	Open	Absent
21	JC	60.6788	133.0105	831	1388.6	-2	7.3	2046.9	14.7	*Pinus contorta* var. *latifolia*	0.708	Needle-Leaf	Open	Absent
22	JC	60.6746	133.0082	850	1256.6	5	23.0	1681.0	17.1	*Picea mariana*	0.662	Needle-Leaf	Closed	Present
23	JC	60.6083	133.0530	1147	1411.3	-24	20.1	168.0	100.0	*Betula glandulosa*	0.288	Bare	Open	Absent
24	JC	60.6086	133.0551	1137	1454.4	-37	13.7	50.0	100.0	Grasses	0.655	Shrub	Closed	Absent
25	JC	60.6082	133.0515	1153	1391.4	-17	10.9	170.0	100.0	*Betula glandulosa*	0.585	Broad-Leaf	Closed	Absent
26	JC	60.4831	133.3096	697	1351.5	-20	4.6	1385.6	14.2	*Populus tremuloides*	0.531	Shrub	Open	Absent
27	JC	60.486	133.2780	756	1378.5	1	9.1	1326.7	28.6	*Pinus contorta* var. *latifolia*	0.487	Broad-Leaf	Closed	Absent

ID	Area	Latitude (° N)	Longitude (° W)	Elevation (m)	PISR (MJ/m^2)	TPI	Organic Mat Thickness (cm)	Vegetation Height (cm)	Canopy Openness (%)	Dominant Species	NDVI	Vegetation Class	Canopy Cover Class	Permafrost
28	SDH	60.5233	128.8468	1063	1397.8	-30	10.4	1284.6	28.6	*Abies lasiocarpa*	0.770	Needle-Leaf	Closed	Absent
29	SDH	60.5251	128.8551	1094	1382.9	-28	11.4	1361.4	27.6	*Abies lasiocarpa*	0.718	Needle-Leaf	Closed	Absent
30	SDH	60.519	128.8663	1188	1398.4	-31	13.1	1368.2	11.6	*Abies lasiocarpa*	0.560	Bare	Open	Absent
31	SDH	60.5148	128.8688	1205	1446.4	-57	14.4	1309.8	18.4	*Abies lasiocarpa*	0.733	Needle-Leaf	Open	Absent
32	SDH	60.5188	128.8813	1493	1361.6	175	1.7	9.0	100.0	*Arctostaphylos rubra*	0.483	Bare	Open	Present
33	SDH	60.5222	128.8716	1296	1376.4	51	7.3	871.8	47.8	*Abies lasiocarpa*	0.827	Needle-Leaf	Closed	Absent
34	SDH	60.5242	128.8689	1228	1366.6	14	11.6	1242.2	35.7	*Abies lasiocarpa*	0.397	Bare	Open	Absent
35	SDH	60.4926	128.8814	1109	1461.4	-39	10.3	1146.3	22.9	*Salix* spp.	0.826	Shrub	Closed	Absent
36	SDH	60.4538	128.9297	972	1427.7	-8	11.8	1657.8	25.5	*Picea glauca*	0.666	Bare	Closed	Absent
37	SDH	60.4619	128.9603	976	1353.0	-61	14.0	1155.0	48.8	*Pinus contorta* var. *latifolia*	0.622	Bare	Closed	Absent
38	SDH	60.4252	129.0074	972	1419.7	19	12.3	780.1	21.9	*Abies lasiocarpa*	0.500	Bare	Open	Absent
39	SDH	60.3049	129.0308	781	1373.3	-12	20.2	1471.2	48.3	*Picea mariana*	0.674	Broad-Leaf	Closed	Absent
40	SDH	60.4304	129.0034	964	1408.6	-9	31.5	230.0	100.0	*Betula glandulosa*	0.698	Needle-Leaf	Closed	Absent
41	SDH	60.5145	128.8694	1208	1441.2	-58	16.9	1708.4	18.4	*Picea glauca*	0.647	Needle-Leaf	Closed	Absent
42	SDH	60.5155	128.8719	1277	1469.1	-4	5.6	406.4	100.0	*Alnus crispa*	0.781	Needle-Leaf	Closed	Absent
43	SDH	60.516	128.8761	1380	1495.3	75	5.2	655.7	38.8	*Abies lasiocarpa*	0.787	Needle-Leaf	Closed	Present
44	SDH	60.5169	128.8775	1437	1477.4	127	3.0	18.0	100.0	*Empetrum nigrum*	0.652	Shrub	Open	Absent
45	SDH	60.5222	128.8753	1345	1387.2	73	6.1	902.0	56.1	*Abies lasiocarpa*	0.393	Bare	Open	Absent
46	SDH	60.5279	128.9186	1442	1467.4	10	10.1	125.0	100.0	Grasses	0.727	Shrub	Open	Absent
47	SDH	60.5277	128.9159	1463	1519.3	19	7.0	19.0	100.0	Moss	0.714	Shrub	Open	Absent
48	SDH	60.531	128.9187	1501	1350.1	86	6.2	15.0	100.0	Lichen	0.505	Tundra	Open	Absent
49	SDH	60.5322	128.9264	1444	1382.4	87	7.7	105.0	100.0	*Betula glandulosa*	0.657	Shrub	Open	Absent
50	SDH	60.5256	128.9090	1532	1522.6	82	2.6	7.0	100.0	*Salix reticulata*	0.569	Tundra	Open	Present
51	SDH	60.5451	128.9064	1438	1237.6	88	6.7	55.0	100.0	Lichen	0.623	Shrub	Open	Absent
52	SDH	60.5441	128.9084	1459	1419.5	111	8.0	19.0	100.0	Lichen	0.627	Shrub	Open	Absent
53	SDH	60.5434	128.9022	1524	1441.1	162	5.3	8.0	100.0	Lichen	0.520	Tundra	Open	Absent
54	SDH	60.5359	128.9056	1415	1504.8	13	5.0	65.0	100.0	*Empetrum nigrum*	0.519	Shrub	Open	Absent
55	SDH	60.4548	128.9333	970	1414.4	-14	10.3	1207.0	28.0	*Populus tremuloides*	0.674	Broad-Leaf	Closed	Absent
56	SDH	60.416	129.0012	950	1409.6	5	14.4	1264.0	21.7	*Pinus contorta* var. *latifolia*	0.681	Broad-Leaf	Closed	Absent
57	SDH	60.3899	129.0498	794	1374.8	0	24.6	917.3	55.7	*Picea mariana*	0.481	Bare	Open	Present

ID	Area	Latitude (° N)	Longitude (° W)	Elevation (m)	PISR (MJ/m^2)	TPI	Organic Mat Thickness (cm)	Vegetation Height (cm)	Canopy Openness (%)	Dominant Species	NDVI	Vegetation Class	Canopy Cover Class	Permafrost
58	F	62.2415	133.3097	1050	1379.0	78	10.2	353.8	100.0	*Picea mariana*	0.659	Shrub	Closed	Present
59	F	62.2734	133.1780	1548	1504.9	96	4.3	13.0	100.0	Lichen	0.417	Tundra	Open	Absent
60	F	62.2697	133.1808	1405	1497.4	10	6.1	210.0	100.0	*Betula glandulosa*	0.591	Shrub	Closed	Absent
61	F	62.2654	133.1829	1342	1446.4	10	13.3	375.1	100.0	*Betula glandulosa*	0.604	Shrub	Closed	Absent
62	F	62.2262	133.3815	633	1285.5	-17	24.8	1155.0	50.8	*Picea mariana*	0.592	Broad-Leaf	Closed	Present
63	F	62.2799	133.2677	1198	1405.1	-35	27.0	684.6	100.0	*Salix* spp.	0.512	Needle-Leaf	Closed	Present
64	F	62.3034	133.3146	1156	1363.2	-58	14.0	180.0	100.0	*Salix* spp.	0.547	Shrub	Closed	Absent
65	F	62.302	133.3186	1192	1312.5	-40	26.7	367.4	100.0	*Picea mariana*	0.546	Needle-Leaf	Closed	Present
66	F	62.2697	133.3128	1084	1403.6	-24	13.9	853.2	44.8	*Picea glauca*	0.428	Needle-Leaf	Closed	Present
67	F	62.2707	133.3092	1157	1424.2	22	7.9	375.1	100.0	*Picea glauca*	0.555	Needle-Leaf	Closed	Present
68	F	62.2678	133.3126	1041	1392.5	-45	17.0	120.0	100.0	*Salix* spp.	0.553	Shrub	Open	Absent
69	F	62.2654	133.3102	1083	1288.9	12	34.0	195.0	100.0	*Picea glauca*	0.542	Shrub	Closed	Present
70	F	62.2642	133.3108	1092	1386.6	36	16.8	405.3	100.0	*Picea mariana*	0.621	Needle-Leaf	Closed	Absent
71	F	62.2745	133.1789	1559	1473.9	86	3.2	4.0	100.0	Lichen	0.275	Tundra	Open	Absent
72	F	62.2481	133.1475	1220	1422.6	10	11.4	1031.0	39.7	*Pinus contorta* var. *latifolia*	0.695	Shrub	Closed	Absent
73	F	62.2461	133.1422	1200	1412.3	3	30.3	110.0	100.0	*Betula glandulosa*	0.550	Shrub	Open	Absent
74	F	62.2139	133.1338	917	1386.5	11	5.2	736.3	24.2	*Salix* spp.	0.644	Shrub	Closed	Absent
75	F	62.2003	133.2469	762	1327.0	3	26.9	736.3	100.0	*Picea mariana*	0.446	Needle-Leaf	Closed	Present
76	F	62.1742	133.1652	659	1296.3	-1	25.6	2088.2	22.7	*Picea glauca*	0.591	Broad-Leaf	Closed	Absent
77	F	62.1851	133.2048	659	1303.8	-13	9.4	1107.3	23.9	*Populus tremuloides*	0.639	Shrub	Closed	Present
78	F	62.2349	133.3677	719	1288.0	5	8.8	311.3	100.0	*Picea mariana*	0.274	Bare	Open	Present
79	F	62.2479	133.3210	898	1272.8	-60	16.3	465.0	100.0	*Picea mariana*	0.333	Bare	Open	Present
80	F	62.0682	132.8821	752	1306.2	-10	23.9	1560.0	29.1	*Picea glauca*	0.469	Broad-Leaf	Closed	Present
81	F	62.0778	132.9035	765	1318.6	2	7.1	1284.6	19.5	*Populus tremuloides*	0.604	Shrub	Closed	Absent
82	F	62.0985	132.9778	790	1342.0	-20	19.0	1211.2	25.2	*Picea mariana*	0.414	Needle-Leaf	Closed	Present
83	F	62.1479	133.1547	759	1334.8	-30	7.1	1560.0	21.4	*Populus tremuloides*	0.574	Needle-Leaf	Closed	Absent
84	F	62.1611	133.2323	754	1333.8	13	17.2	292.3	100.0	*Salix* spp.	0.560	Shrub	Open	Present

ID	Area	Latitude (° N)	Longitude (° W)	Elevation (m)	PISR (MJ/m^2)	TPI	Organic Mat Thickness (cm)	Vegetation Height (cm)	Canopy Openness (%)	Dominant Species	NDVI	Vegetation Class	Canopy Cover Class	Permafrost
85	D	64.0826	139.5916	1010	1349.2	101	14.0	331.3	100.0	*Betula glandulosa*	0.636	Needle-Leaf	Open	Present
86	D	64.165	139.8029	1224	1403.5	102	5.1	26.0	100.0	*Betula glandulosa*	0.560	Tundra	Open	Present
87	D	64.0836	139.4429	320	1191.9	-39	33.0	338.2	100.0	*Picea mariana*	0.611	Needle-Leaf	Open	Present
88	D	64.0765	139.4807	480	1221.4	-38	33.0	215.0	100.0	*Picea mariana*	0.605	Shrub	Open	Present
89	D	64.0767	139.4782	482	1203.2	-33	17.8	1102.7	33.6	*Picea glauca*	0.625	Broad-Leaf	Closed	Absent
90	D	64.0721	139.4863	559	1239.0	0	26.3	641.7	45.4	*Picea mariana*	0.600	Needle-Leaf	Closed	Present
91	D	64.0585	139.4695	680	1295.4	149	14.4	780.1	37.7	*Populus tremuloides*	0.681	Needle-Leaf	Open	Absent
92	D	64.1274	139.6991	1125	1218.9	66	15.3	90.0	100.0	*Betula glandulosa*	0.647	Shrub	Open	Absent
93	D	64.0525	139.4688	467	1286.8	16	9.9	1088.3	51.7	*Populus tremuloides*	0.680	Broad-Leaf	Open	Absent
94	D	64.0531	139.4755	494	1276.3	19	13.9	1478.1	100.0	*Populus tremuloides*	0.662	Shrub	Open	Absent
95	D	64.0887	139.5509	853	1334.1	30	10.8	715.6	100.0	*Picea mariana*	0.676	Shrub	Open	Absent
96	D	64.0894	139.5620	852	1259.5	24	11.0	609.8	100.0	*Picea mariana*	0.558	Needle-Leaf	Open	Absent
97	D	64.1597	139.8016	1120	1425.7	49	21.7	150.0	100.0	*Betula glandulosa*	0.757	Shrub	Open	Absent
98	D	64.1674	139.8029	1223	1400.8	103	4.2	20.0	100.0	*Betula glandulosa*	0.540	Tundra	Open	Present
99	D	64.1624	139.8081	1181	1406.0	78	8.9	30.0	100.0	*Betula glandulosa*	0.591	Shrub	Open	Present
100	D	64.1275	139.6958	1130	1139.3	76	13.4	110.0	100.0	*Betula glandulosa*	0.694	Shrub	Open	Absent
101	D	64.0717	139.4452	402	1225.0	24	10.1	1568.8	19.8	*Picea mariana*	0.659	Broad-Leaf	Closed	Present
102	D	64.1884	140.0061	1169	1286.6	78	11.7	85.0	100.0	*Betula glandulosa*	0.662	Shrub	Open	Present
103	D	64.1873	139.9956	1231	1401.8	133	2.2	8.0	100.0	*Betula glandulosa*	0.595	Tundra	Open	Present
104	D	64.1901	139.9999	1226	1303.7	130	8.0	15.0	100.0	*Betula glandulosa*	0.490	Tundra	Open	Present
105	D	64.1876	140.0142	1084	1295.4	15	13.4	200.0	100.0	*Betula glandulosa*	0.653	Shrub	Open	Absent
106	D	64.077	139.4567	424	1219.3	10	26.1	286.2	100.0	*Picea mariana*	0.573	Needle-Leaf	Open	Present
107	D	64.0737	139.4640	459	1226.3	4	6.7	1063.4	27.7	*Populus tremuloides*	0.571	Needle-Leaf	Open	Absent
108	D	64.0704	139.4661	458	1238.1	-24	26.4	485.1	38.0	*Picea mariana*	0.512	Bare	Open	Present
109	D	64.0579	139.4767	596	1269.7	49	16.0	1326.7	28.4	*Picea glauca*	0.592	Bare	Open	Absent
110	D	64.0677	139.3956	878	1284.2	153	1.0	7.0	100.0	Grasses	0.152	Bare	Open	Absent
111	D	64.0692	139.3905	878	1255.6	124	5.1	331.3	100.0	Grasses	0.519	Tundra	Open	Absent
112	D	64.0699	139.3852	865	1308.8	105	7.8	1031.0	19.0	*Betula papyrifera*	0.733	Shrub	Open	Absent
113	D	64.0649	139.3790	731	1339.7	7	17.7	1326.7	22.2	*Picea glauca*	0.319	Bare	Closed	Absent
114	D	64.0818	139.4433	324	1192.2	-38	23.2	511.4	40.6	*Picea mariana*	0.728	Broad-Leaf	Closed	Present

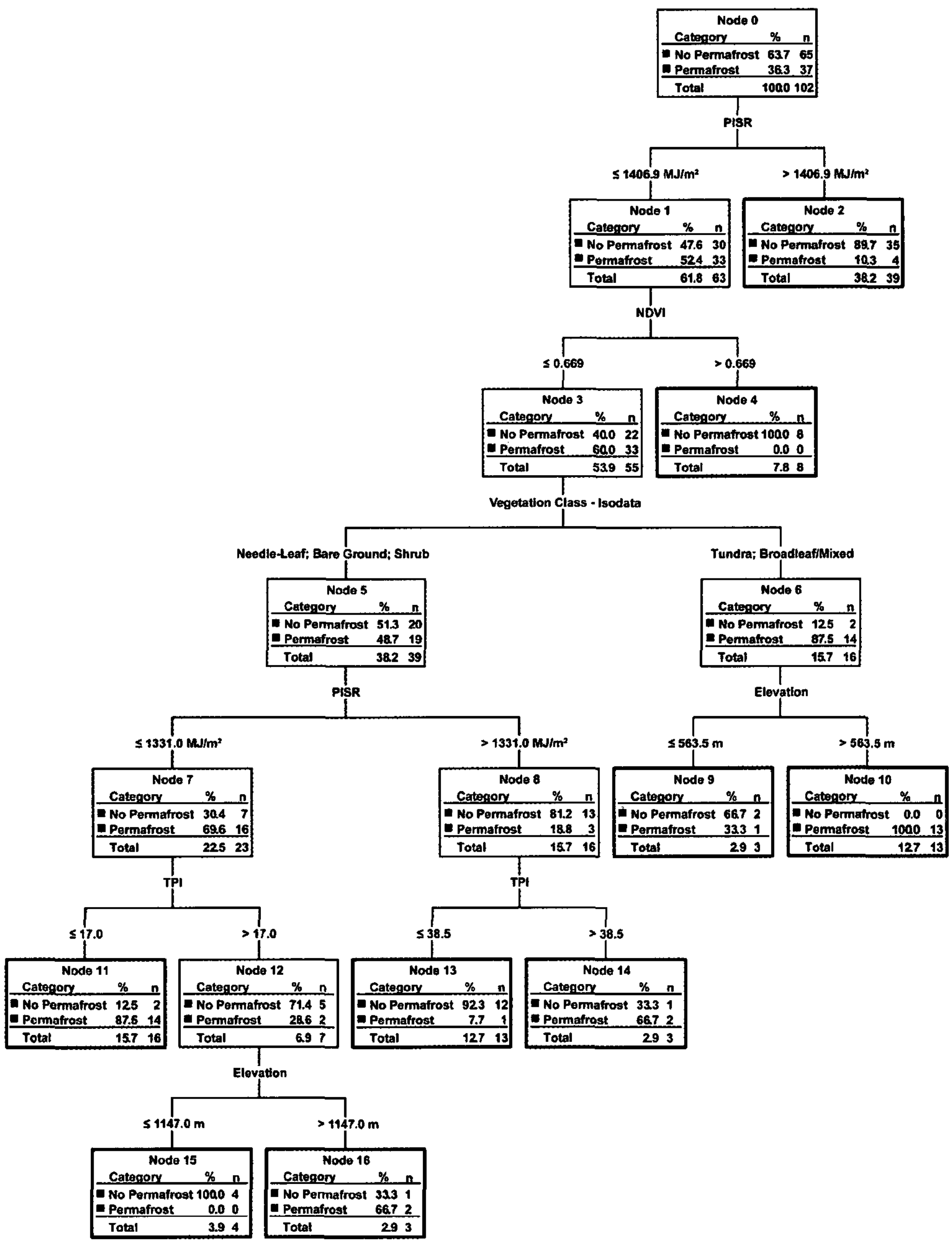
ID	Area	Latitude (° N)	Longitude (° W)	Elevation (m)	PISR (MJ/m^2)	TPI	Organic Mat Thickness (cm)	Vegetation Height (cm)	Canopy Openness (%)	Dominant Species	NDVI	Vegetation Class	Canopy Cover Class	Permafrost
115	K	63.9309	135.2433	1469	1381.2	41	10.6	110.0	100.0	*Betula glandulosa*	0.389	Bare	Open	Present
116	K	63.926	135.2347	1444	1476.5	52	8.8	115.0	100.0	*Betula glandulos*	0.450	Bare	Open	Absent
117	K	63.9184	135.2342	1236	1442.5	16	11.1	235.0	100.0	*Betula glandulosa*	0.541	Broad-Leaf	Closed	Absent
118	K	63.9257	135.2260	1316	1242.6	-66	11.7	455.6	100.0	*Abies lasiocarpa*	0.357	Bare	Open	Absent
119	K	63.914	135.2323	1044	1362.3	-133	21.7	180.0	100.0	*Salix* spp.	0.500	Shrub	Closed	Present
120	K	63.9146	135.2320	1063	1374.1	-116	9.6	465.0	100.0	*Salix* spp.	0.512	Shrub	Closed	Absent
121	K	63.9516	135.2868	1242	1366.7	9	6.4	80.0	100.0	*Betula glandulosa*	0.428	Bare	Open	Absent
122	K	63.9128	135.2307	1078	1143.6	-104	22.6	175.0	100.0	*Alnus crispa*	0.500	Shrub	Closed	Present
123	K	63.9097	135.2244	1230	1178.4	0	25.9	609.8	100.0	*Abies lasiocarpa*	0.538	Shrub	Closed	Present
124	K	63.9488	135.4127	730	1239.3	10	12.8	886.6	37.4	*Betula papyrifera*	0.403	Broad-Leaf	Open	Present
125	K	63.955	135.4285	678	1272.1	-2	20.6	331.3	100.0	*Picea mariana*	0.500	Needle-Leaf	Closed	Present
126	K	63.9427	135.1920	1803	1382.9	129	2.3	12.0	100.0	Grasses	0.176	Tundra	Open	Present
127	K	63.9428	135.2051	1721	1137.6	92	8.3	22.0	100.0	*Salix* spp.	0.419	Tundra	Open	Present
128	K	63.926	135.2316	1434	1466.6	40	5.8	19.0	100.0	*Betula glandulosa*	0.384	Tundra	Open	Absent
129	K	63.9247	135.2350	1420	1466.3	56	5.4	65.0	100.0	*Betula glandulosa*	0.400	Broad-Leaf	Open	Absent
130	K	63.9027	135.5243	801	1261.6	-9	24.1	399.4	100.0	*Picea mariana*	0.600	Shrub	Closed	Present
131	K	63.9187	135.3244	871	1283.9	-28	32.0	445.1	100.0	*Picea glauca*	0.488	Broad-Leaf	Open	Present
132	K	63.8866	135.2717	1418	1442.7	85	7.1	22.0	100.0	*Betula glandulosa*	0.418	Broad-Leaf	Open	Absent
133	K	63.8977	135.2672	1397	1384.4	89	17.3	140.0	100.0	*Betula glandulosa*	0.476	Broad-Leaf	Open	Absent
134	K	63.9419	135.2174	1693	1445.1	87	3.2	10.0	100.0	Grasses	0.354	Tundra	Open	Present
135	K	63.9224	135.5244	695	1293.3	-2	16.6	1022.8	39.8	*Picea mariana*	0.405	Shrub	Closed	Present
136	K	63.9086	135.3036	944	1319.2	-21	16.2	989.2	39.9	*Salix* spp.	0.272	Bare	Open	Present

APPENDIX C

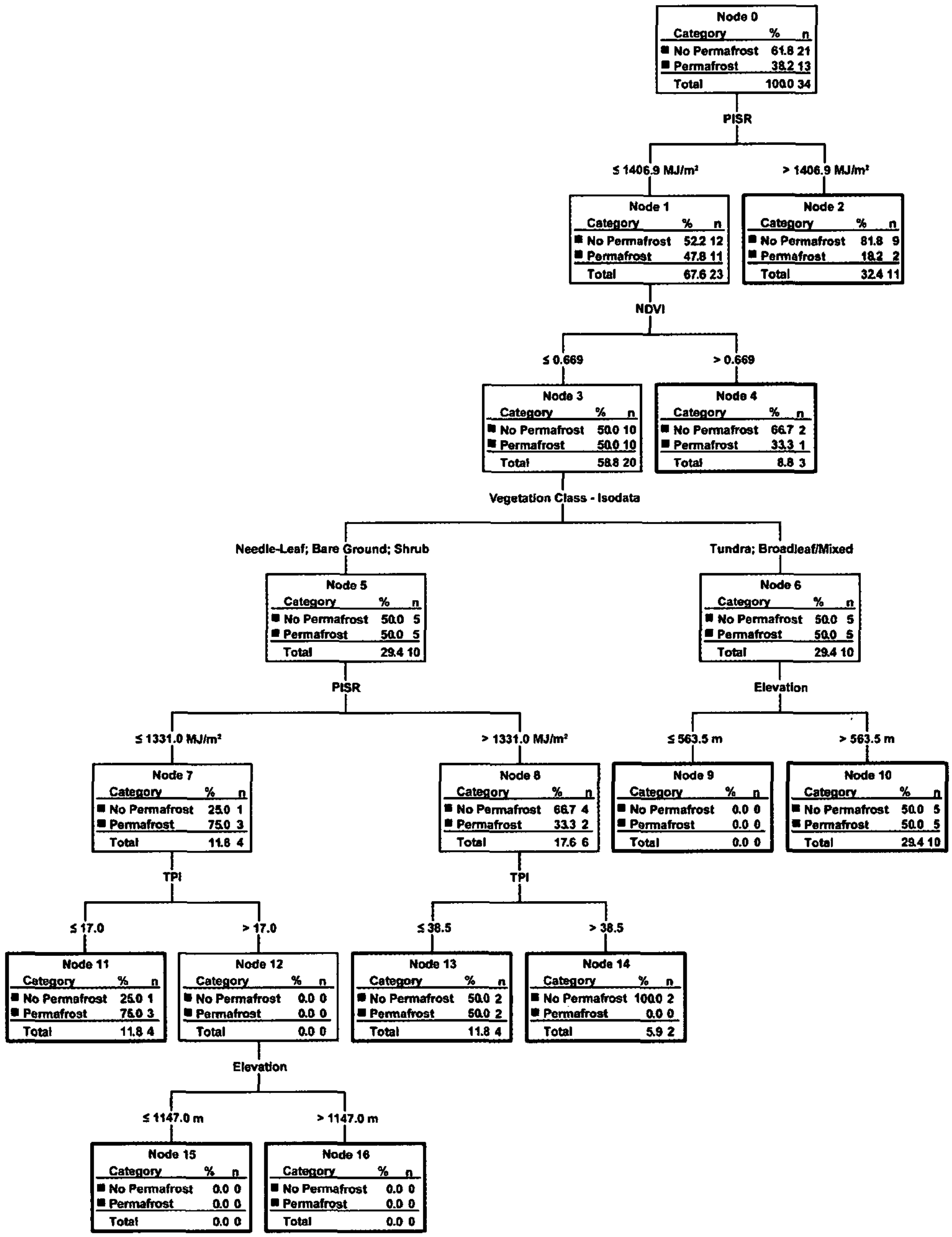
Regression trees comparing K-Means and ISOData. Models use five variables: Elevation, PISR, TPI, NDVI, Vegetation Classification (K-Means or ISOData).

Accuracy of training and testing trees for ISOData and K-Means classification.

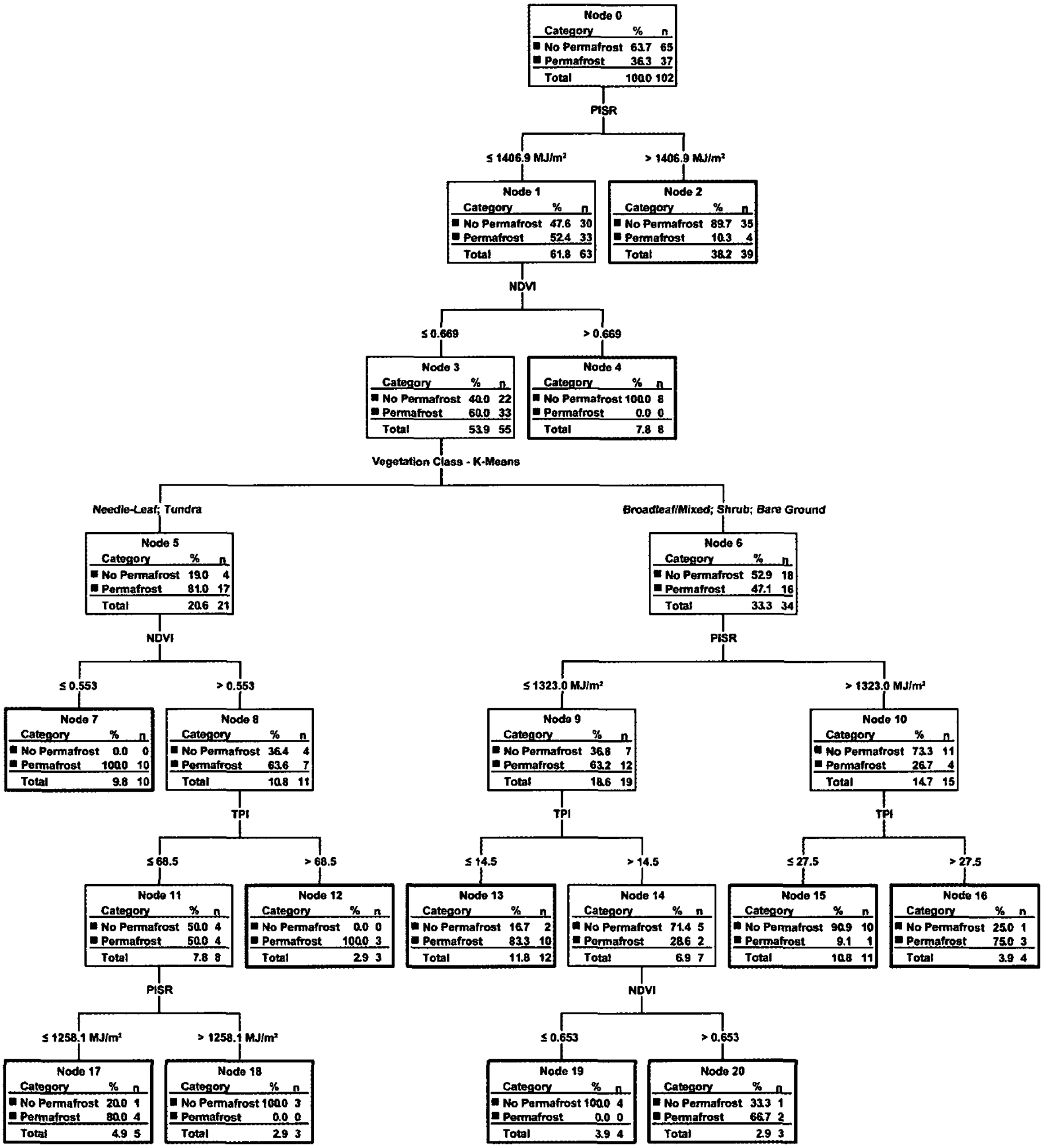
	Training Data		Testing Data	
	ISOData	K-Means	ISOData	K-Means
Accuracy	90.2%	90.2%	61.8%	**64.7%**



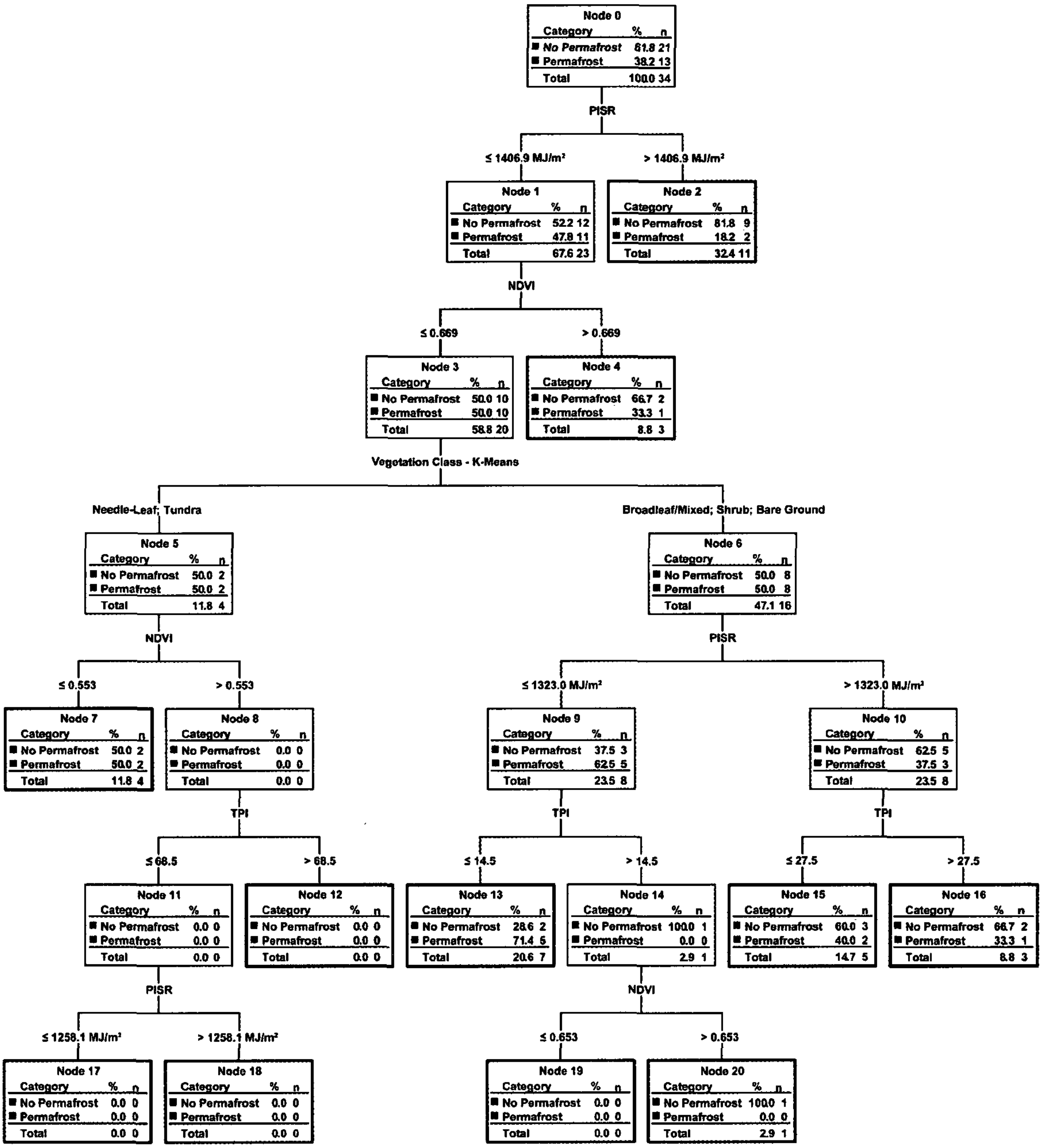
Training tree using ISOData-based classification



Testing tree using ISOData-based classification



Training tree using K-Means-based classification



Testing tree using K-Means-based classification

APPENDIX D

Accuracies of training and testing regression trees for all sites and for each of the three test sample groups.

DEM-Based tree

	Training				Testing			
	All Sites	Sample 1	Sample 2	Sample 3	All Sites	Sample 1	Sample 2	Sample 3
Accuracy	86.0 %	81.4 %	86.3 %	86.3 %	N/A	76.5 %	73.5 %	79.4 %

Field-Based tree

	Training				Testing			
	All Sites	Sample 1	Sample 2	Sample 3	All Sites	Sample 1	Sample 2	Sample 3
Accuracy	77.9 %	86.3 %	84.3 %	86.3 %	N/A	76.5 %	76.5 %	76.5 %

Field and DEM-Based tree

	Training				Testing			
	All Sites	Sample 1	Sample 2	Sample 3	All Sites	Sample 1	Sample 2	Sample 3
Accuracy	83.1 %	84.3 %	82.4 %	84.3 %	N/A	61.8 %	79.4 %	70.6 %

Imagery-Based tree

	Training				Testing			
	All Sites	Sample 1	Sample 2	Sample 3	All Sites	Sample 1	Sample 2	Sample 3
Accuracy	69.1 %	70.6 %	76.5 %	74.5 %	N/A	58.8 %	41.2 %	55.9 %

Imagery and DEM-Based tree

	Training				Testing			
	All Sites	Sample 1	Sample 2	Sample 3	All Sites	Sample 1	Sample 2	Sample 3
Accuracy	81.6 %	79.4 %	84.3 %	86.3 %	N/A	73.5 %	55.9 %	67.6 %

APPENDIX E

User's accuracies for each model for individual study areas.

Dawson

	Permafrost	
	Improbable	Probable
DEM-Based	93.8%	92.9%
Field-Based	88.9%	100.0%
Field and DEM-Based	100.0%	100.0%
Imagery-Based	85.7%	56.5%
Imagery and DEM-Based	93.8%	87.5%

Faro

	Permafrost	
	Improbable	Probable
DEM-Based	92.3%	92.9%
Field-Based	78.6%	84.6%
Field and DEM-Based	N/A	N/A
Imagery-Based	68.8%	81.8%
Imagery and DEM-Based	100.0%	77.8%

Keno

	Permafrost	
	Improbable	Probable
DEM-Based	100.0%	92.9%
Field-Based	80.0%	91.7%
Field and DEM-Based	N/A	N/A
Imagery-Based	83.3%	75.0%
Imagery and DEM-Based	100.0%	92.9%

Johnson's Crossing

	Permafrost	
	Improbable	Probable
DEM-Based	91.3%	75.0%
Field-Based	91.7%	33.3%
Field and DEM-Based	N/A	N/A
Imagery-Based	81.5%	N/A
Imagery and DEM-Based	100.0%	71.4%

Sa Dena Hes

	Permafrost	
	Improbable	Probable
DEM-Based	86.7%	N/A
Field-Based	96.2%	75.0%
Field and DEM-Based	N/A	N/A
Imagery-Based	86.7%	N/A
Imagery and DEM-Based	86.7%	N/A

User's accuracies for each model created using all data.

DEM-Based

	Permafrost	
	Improbable	Probable
Training Model	89.6%	85.7%
Testing Model	80.0%	64.3%

Field-Based

	Permafrost	
	Improbable	Probable
Training Model	81.0%	95.7%
Testing Model	72.0%	66.7%

Field and DEM-Based

	Permafrost	
	Improbable	Probable
Training Model	86.5%	96.4%
Testing Model	74.1%	85.7%

Imagery-Based

	Permafrost	
	Improbable	Probable
Training Model	73.5%	78.9%
Testing Model	61.5%	37.5%

Imagery and DEM-Based

	Permafrost	
	Improbable	Probable
Training Model	92.3%	86.5%
Testing Model	73.7%	53.3%