CONTRIBUTIONS OF OPTICAL REMOTE SENSING TO PERMAFROST MAPPING IN DONNELLY TRAINING AREA, ALASKA

Kiran Thapa
Southern Illinois University Carbondale, thapakiran668@gmail.com

Follow this and additional works at: https://opensiuc.lib.siu.edu/theses

Recommended Citation
Thapa, Kiran, "CONTRIBUTIONS OF OPTICAL REMOTE SENSING TO PERMAFROST MAPPING IN DONNELLY TRAINING AREA, ALASKA" (2020). Theses. 2764.
https://opensiuc.lib.siu.edu/theses/2764

This Open Access Thesis is brought to you for free and open access by the Theses and Dissertations at OpenSIUC. It has been accepted for inclusion in Theses by an authorized administrator of OpenSIUC. For more information, please contact opensiuc@lib.siu.edu.
CONTRIBUTIONS OF OPTICAL REMOTE SENSING TO PERMAFROST MAPPING IN
DONNELLY TRAINING AREA, ALASKA

by

Kiran Thapa

B.S., Tribhuvan University Kathmandu, Nepal, 2008

A Thesis
Submitted in Partial Fulfillment of the Requirements for the
Master of Science Degree

Department of Geography and Environmental Resources
in the Graduate School
Southern Illinois University Carbondale
August 2020
AN ABSTRACT OF THE THESIS OF

Kiran Thapa, for the Master of Science degree in Geography and Environmental Resources, presented on April 8, 2020, at Southern Illinois University Carbondale.

TITLE: CONTRIBUTIONS OF OPTICAL REMOTE SENSING TO PERMAFROST MAPPING IN DONNELLY TRAINING AREA, ALASKA

MAJOR PROFESSOR: Dr. Guangxing Wang

Permafrost occupies about a quarter of the northern hemisphere land with 25.5 million ha. Global warming and anthropogenic activities affect the dynamics of permafrost. Snow and permafrost, in turn, serve as an indicator of climate change and human activity disturbance. The dynamics of permafrost are often estimated using interferometric Synthetic Aperture Radar (InSAR) methods. However, acquiring and processing InSAR images is costly and computation intensive. Due to various spectral variables and indices available from optical images, Landsat satellite images that are free-downloadable provide the potential for studying and monitoring changes of permafrost.

The overall objective of this study was to explore the use of optical images as a cost-effective method to map permafrost in Donnelly Training Area (DTA) - an installation located in Alaska. First, Landsat 8 OLI/TIRS images from January 2014 to December 2018 were used to calculate various remote sensing variables. The variables included Land Surface Temperature (LST), albedo, Soil Moisture index (SMI), Normalized Difference Vegetation Index (NDVI), Normalized Difference Snow Index (NDSI), Normalized Difference Built-up Index (NDBI), Normalized Difference Water index (NDWI), Simple Ratio (SR), Soil Adjusted Vegetation Index (SAVI), Normalized Burn Ratio (NBR), Triangular Vegetation Index (TVI), Visible Atmospherically Resistant Index (VARI), and Active Layer Thickness (ALT). Moreover, elevation, slope, and aspect were obtained from a digital elevation model (DEM). The variables
were used to estimate the probabilities of permafrost presence (POP) for DTA. The logistic and linear models were respectively selected and optimized based on logistic and linear stepwise regression for the estimation of and ALT. A total of 414 field observations that were collected from 1994 to 2012 were utilized for validation of models.

The results showed that the POP in DTA was significantly affected by all the factors except aspect and EVI. The factor that was most correlated with \( \ln((1-\text{POP})/\text{POP}) \) was elevation, then NDVI, albedo, ALT, LST, NDWI, NDSI, slope, TVI, RSR, SMI, NDBI, SR, SAVI, NBR and VARI. A total of six prediction models were obtained. The elevation, NDVI, LST, TVI, ALT, SLOPE, RSR, SMI, NBR, and NDSI were finally chosen in the best model 5.6 with the smallest relative root mean square error (RMSE) and Akaike information criterion (AIC). The albedo used in previous studies was excluded in the final model, implying that the albedo was not critical to the prediction of POP. In addition to the previously used elevation, NDVI and SMI, other predictors including LST, TVI, ALT, SLOPE, RSR, NBR and NDSI could not be ignored in the prediction of POP. The model generated reasonable spatial distribution of POP in which POP had greater values in the east, northeast, north and northwest parts and smaller in the south and southwest parts. Except for NDVI, NDWI, NDSI, aspect and RSR, moreover, all other predictors showed significant contributions to prediction of ALT. The SMI, ELEVATION, SAVI, NDBI, SLOPE, LST, SR, EVI, VARI and TVI were finally selected in the best model 5.14 with the smallest relative RMSE and AIC. The ALT highly varied over the study area with the spatial patterns inversely consistent with those of POP.

The results are essential for the governments, policymakers, and other concerned stakeholders to estimate the degradation of permafrost in DTA and minimize the risk of policy decision-making for land use management and planning. This study will help to understand the
global climate change, changing ecosystems, increasing concentration in the atmosphere, and human activity induced disturbance.
ACKNOWLEDGMENTS

I would like to express my gratitude to my advisor, Professor Dr. Guangxing Wang. He has been a great mentor for me throughout my graduate study in Southern Illinois University (SIU). I would like to thank my committee members Dr. Justin Schoof and Dr. Rupou Li for their valuable comments and great effort to improve my thesis. I would also like to thank all other professors and staff at the Department of Geography and Environmental Resources for their support. This research would be incomplete without funding and I would like to thank the U.S. Army, Construction Engineering Research Laboratory (CERL), Engineer Research and Development Center (ERDC).
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>i</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xiii</td>
</tr>
<tr>
<td>CHAPTERS</td>
<td></td>
</tr>
<tr>
<td>CHAPTER 1 – INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Research Statement</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Objectives and Research Questions</td>
<td>5</td>
</tr>
<tr>
<td>1.4 Significance of the Study</td>
<td>5</td>
</tr>
<tr>
<td>CHAPTER 2 – LITERATURE REVIEW</td>
<td>6</td>
</tr>
<tr>
<td>2.1 Permafrost Distribution in Northern High Latitudes and its Degradation</td>
<td>6</td>
</tr>
<tr>
<td>2.2 Mapping Methods for Permafrost</td>
<td>10</td>
</tr>
<tr>
<td>2.3 Remote Sensing Imagery-Derived Vegetation Indices and Factors Affecting Permafrost Distribution</td>
<td>12</td>
</tr>
<tr>
<td>2.3.1 Normalized Difference Vegetation Index (NDVI)</td>
<td>12</td>
</tr>
<tr>
<td>2.3.2 Albedo</td>
<td>12</td>
</tr>
<tr>
<td>2.3.3 Land Surface Temperature (LST)</td>
<td>13</td>
</tr>
<tr>
<td>2.3.4 Soil Moisture Index (SMI)</td>
<td>14</td>
</tr>
<tr>
<td>2.3.5 Normalized Difference Snow Index (NDSI)</td>
<td>14</td>
</tr>
<tr>
<td>2.3.6 Normalized Difference Water Index (NDWI)</td>
<td>15</td>
</tr>
</tbody>
</table>
4.2.11 Simple Ratio (SR) ................................................................. 37
4.2.12 Soil Adjusted Vegetation Index (SAVI) .................................. 38
4.2.13 Enhanced Vegetation Index (EVI) ........................................... 38
4.2.14 Normalized Burn Ration (NBR) .............................................. 39
4.2.15 Triangular Vegetation Index (TVI) .......................................... 39
4.2.16 Reduced Simple Ratio (RSR) .................................................. 39
4.2.17 Visible Atmospherically Resistant Index (VARI) ...................... 40
4.2.18 Active Layer Thickness (ALT) ................................................. 40
4.3 Correlation Analysis and Selection of Model Factors ....................... 42
4.4 Permafrost Distribution Mapping of DTA ..................................... 43
4.5 Validations of Models ............................................................... 43

CHAPTER 5 – RESULTS ...................................................................... 45
5.1 Optical Remote Sensing Factors .................................................. 45
5.2 Contribution of Factors ............................................................. 63
  5.2.1 Correlation Analysis .................................................................. 64
5.3 Mapping probability of permafrost presence ................................ 72
5.4 Mapping of Active Layer Thickness ............................................. 82
5.5 Validation ................................................................................. 91

CHAPTER 6 – DISCUSSION ............................................................... 95
6.1 Performance of logistic model with original predictors .................. 95
6.2 Contribution of other predictors for predicting probability of permafrost .... 96
6.3 Mapping active layer thickness ................................................... 97
6.4 Spatial distribution of Permafrost and active layer thickness .............. 98
6.5 Limitation ................................................................................................................. 98

CHAPTER 7 – CONCLUSIONS ......................................................................................... 100

REFERENCES .................................................................................................................. 102

APPENDICES

APPENDIX A – PERMAFROST DISTRIBUTION IN NORTHERN HEMISPHERE 122

APPENDIX B – LAND SURFACE BRIGHTNESS TEMPERATURE ......................... 123

VITA ................................................................................................................................. 124
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 3.1 - Landsat datasets were used with different periods</td>
<td>21</td>
</tr>
<tr>
<td>Table 3.2 - Summary of band designations and pixel size for Landsat 8 Satellite Sensor</td>
<td>22</td>
</tr>
<tr>
<td>Table 3.3 - Aspect values from Map</td>
<td>26</td>
</tr>
<tr>
<td>Table 5.1 - Burning classes based on Normalized Burn Ratio (NBR)</td>
<td>56</td>
</tr>
<tr>
<td>Table 5.2 - Correlation summary of Ln((1-POP/POP)) and ALT with the remote sensing image derived factors and among the factors</td>
<td>65</td>
</tr>
<tr>
<td>Table 5.3 - Interpreting the size of a Correlation Coefficient</td>
<td>71</td>
</tr>
<tr>
<td>Table 5.4 - Summary of models obtained for prediction of probability of permafrost presence using stepwise regression</td>
<td>72</td>
</tr>
<tr>
<td>Table 5.5 - The coefficients and variation inflation factors (VIF) of the models obtained for prediction of probability of permafrost presence using stepwise regression</td>
<td>73</td>
</tr>
<tr>
<td>Table 5.6 - Summary of models obtained for prediction of active layer thickness using stepwise regression with VIF equal to 10</td>
<td>83</td>
</tr>
<tr>
<td>Table 5.7 - The regression coefficients and VIF values of the models for prediction of active layer thickness using stepwise regression</td>
<td>84</td>
</tr>
<tr>
<td>Table 5.8 - Relative RMSE of POP and ALT estimates from the developed models</td>
<td>92</td>
</tr>
<tr>
<td>Table 5.9 - Akaike Information Criterion (AIC) for the models of POP in DTA</td>
<td>93</td>
</tr>
<tr>
<td>Table 5.10 - Akaike Information Criterion (AIC) for Active Layer Thickness (ALT) in DTA</td>
<td>94</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 3.1 - Location map of the study area shown by a Landsat 8 OLI (color composite red, green, and blue: 4, 3, and 2).</td>
<td>18</td>
</tr>
<tr>
<td>Figure 3.2 - Land Cover in Donnelly Training Area</td>
<td>19</td>
</tr>
<tr>
<td>Figure 3.3 - Average Monthly Temperature in Fairbanks, Alaska from 201-2018</td>
<td>20</td>
</tr>
<tr>
<td>Figure 3.4 - Monthly precipitations in Fairbanks, Alaska from 2014-2018</td>
<td>20</td>
</tr>
<tr>
<td>Figure 3.5 - Landsat 8 images acquired in different years (2014 to 2018) (color composite red, green, and blue: 4, 3, and 2). a) Color composite image of January 30th, 2014, b) Color composite image of April 23rd, 2015, c) Color composite image of March 24th, 2016, d) Color composite image of January 22nd, 2017, and e) Color composite image of February 10th, 2018</td>
<td>23</td>
</tr>
<tr>
<td>Figure 3.6 - Geological rock units in the study area (top) and Elevation map of DTA (bottom).</td>
<td>24</td>
</tr>
<tr>
<td>Figure 3.7 - Aspect map of the DTA (top) and Slope map of the DTA area (bottom)</td>
<td>25</td>
</tr>
<tr>
<td>Figure 3.8 - DTA overlain with presence-absence of near surface (within 1m) permafrost field observations and a Landsat 8 OLI (false color composite [red, green, and blue: 6, 5, and 4]).</td>
<td>27</td>
</tr>
<tr>
<td>Figure 3.9 - Predicted Active Layer Thickness of 2010 to 2019 and developed by Geophysical Institute Permafrost Lab</td>
<td>28</td>
</tr>
<tr>
<td>Figure 4.1 - Detail flowchart for the study of permafrost mapping in the Donnelly Training Area</td>
<td>30</td>
</tr>
<tr>
<td>Figure 4.2 - Correction processes and units for pre-processing remotely sensed imagery</td>
<td>33</td>
</tr>
</tbody>
</table>
Figure 4.3 - Showing the ground layer, frozen layer, active layer thickness and situation of Permafrost ................................................................. 41

Figure 5.1 - The average normalized difference vegetation index of DTA................................. 46

Figure 5.2 - Enhanced Vegetation Index (EVI) in the Donnelly Training Area, Alaska. ........ 47

Figure 5.3 - Visible Atmospherically Resistant Index (VARI) in the Donnelly Training Area. . 48

Figure 5.4 - Simple Ratio (SR) in the Donnelly Training Area, Alaska................................. 49

Figure 5.5 - Reduced Simple Ration (RSR) in the Donnelly Training Area, Alaska................. 49

Figure 5.6 - Proportion of vegetation of Donnelly Training Area in 02/10/2018.................... 50

Figure 5.7 - Triangular Vegetation Index (TVI) in the Donnelly Training Area Alaska .......... 51

Figure 5.8 - The Average Soil Moisture Index (SMI) in the Donnelly Training Area. .......... 52

Figure 5.9 - The average Soil Adjusted Vegetation Index (SAVI) in the Donnelly Training Area, Alaska ................................................................. 53

Figure 5.10 - The Average Normalized Difference Water Index (NDWI) in the Donnelly Training Area. ................................................................. 54

Figure 5.11 - The Average Normalized Difference Snow Index in the Donnelly Training Area................................................................. 55

Figure 5.12 - The Average Normalized Burn Ratio (NBR) in the Donnelly Training Area, Alaska ................................................................. 56

Figure 5.13 - The Average Land Surface Emissivity in the Donnelly Training Area. .......... 57

Figure 5.14 - The Average Land Surface Temperature (LST) in the Donnelly Training Area. ... 58

Figure 5.15 - The Average Albedo in the Donnelly Training Area ........................................ 60

Figure 5.16 - The Average Normalized Difference Built-up Index (NDBI) in the Donnell Training Area. ................................................................. 61
Figure 5.17 - The Active Layer Thickness (ALT) in the Donnelly Training Area, Alaska. .......... 62

Figure 5.18 - The distribution probabilities of permafrost presence (POP) in Donnelly Training Area, Alaska, based on the model proposed by Wang (2017)................. 63

Figure 5.19 - The randomly selected 500 points in the DTA area, Alaska.................................. 64

Figure 5.20 - Scatter plot with fit line of Presence of permafrost (POP) with predictor. .......... 66

Figure 5.21 - The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on model 1.................................................................76

Figure 5.22 - The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 2.................................................................77

Figure 5.23 - The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 3.................................................................78

Figure 5.24 - The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 4.................................................................79

Figure 5.25 - The spatial distribution of permafrost probabilities in Donnelly Training Area, ... 80

Figure 5.26 - The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 6.................................................................81

Figure 5.27 - The spatial distributions of Active Layer Thickness estimates obtained by model 5.7 to model 5.14 in the Donnelly Training Area, Alaska. .................90-90
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Full Title</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active Layer Thickness</td>
<td>ALT</td>
</tr>
<tr>
<td>Akaike Information Criterion</td>
<td>AIC</td>
</tr>
<tr>
<td>Brightness Temperature</td>
<td>BT</td>
</tr>
<tr>
<td>Centimeter</td>
<td>CM</td>
</tr>
<tr>
<td>Department of Defense</td>
<td>DoD</td>
</tr>
<tr>
<td>Digital Elevation Model</td>
<td>DEM</td>
</tr>
<tr>
<td>Donnelly Training Area</td>
<td>DTA</td>
</tr>
<tr>
<td>Dry Bareness Index</td>
<td>DBSI</td>
</tr>
<tr>
<td>Dry- Built-Up Index</td>
<td>DBI</td>
</tr>
<tr>
<td>Enhanced Vegetation Index</td>
<td>EVI</td>
</tr>
<tr>
<td>Geographic Information System</td>
<td>GIS</td>
</tr>
<tr>
<td>Inferometric Synthetic Aperture Radar</td>
<td>InSAR</td>
</tr>
<tr>
<td>Land Surface Emissivity</td>
<td>LSE</td>
</tr>
<tr>
<td>Land Surface Temperature</td>
<td>LST</td>
</tr>
<tr>
<td>Land Use and Land Cover</td>
<td>LULC</td>
</tr>
<tr>
<td>Mean Annual Air Temperature</td>
<td>MAAT</td>
</tr>
<tr>
<td>Moderate Resolution Imaging Spectrometer</td>
<td>MODIS</td>
</tr>
<tr>
<td>Near Infrared</td>
<td>NIR</td>
</tr>
<tr>
<td>Normalized Burn Ratio</td>
<td>NBR</td>
</tr>
<tr>
<td>Normalized Difference Vegetation Index</td>
<td>NDVI</td>
</tr>
<tr>
<td>Northern Ecosystem Soil Temperature</td>
<td>NEST</td>
</tr>
<tr>
<td>Term</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Operational Land Imager</td>
<td>OLI</td>
</tr>
<tr>
<td>Probability of Permafrost Presence</td>
<td>POP</td>
</tr>
<tr>
<td>Qinghai-Tibet Plateau</td>
<td>QTP</td>
</tr>
<tr>
<td>Reduced Simple Ration</td>
<td>RSR</td>
</tr>
<tr>
<td>Root Mean Square Error</td>
<td>RMSE</td>
</tr>
<tr>
<td>Synthetic Aperture Radar</td>
<td>SAR</td>
</tr>
<tr>
<td>Shortwave Near Infrared</td>
<td>SWIR</td>
</tr>
<tr>
<td>Simple Ratio</td>
<td>SR</td>
</tr>
<tr>
<td>Soil Adjusted Vegetation Index</td>
<td>SAVI</td>
</tr>
<tr>
<td>Soil Moisture Index</td>
<td>SMI</td>
</tr>
<tr>
<td>Surface Air Temperature</td>
<td>SAT</td>
</tr>
<tr>
<td>Synthetic Aperture Radar</td>
<td>SAR</td>
</tr>
<tr>
<td>Temperature at the Top of the Permafrost</td>
<td>TTOP</td>
</tr>
<tr>
<td>Thermal Infrared Sensors</td>
<td>TIRS</td>
</tr>
<tr>
<td>Top of Atmosphere</td>
<td>TOA</td>
</tr>
<tr>
<td>Triangular Vegetation Index</td>
<td>TVI</td>
</tr>
<tr>
<td>United States Geological Survey</td>
<td>USGS</td>
</tr>
<tr>
<td>Visible Atmospherically Resistant Index</td>
<td>VARI</td>
</tr>
</tbody>
</table>
CHAPTER 1
INTRODUCTION

1.1 Background

Permafrost is frozen soil or rock, including ice or organic materials, which remains at or below 0 °C for at least two or more consecutive years (Woo, 2012; Van Everdingen, 2005). Some permafrost has been frozen for thousands of years. About a quarter of the land in the Northern Hemisphere is underlain by permafrost (Zhang et al., 1999), 25.5 million hectares (including glaciers) or about 23% of the land area. The remaining landforms such as sand wedges, patterned ground pingo scars, and rock glaciers give an indication about where permafrost once existed. According to Ferrians (1965) and Swanson (2017), permafrost is classified into four classes depending on the percentage underlain by permafrost: continuous permafrost-covering more than 90% of the landscapes, discontinuous permafrost-covering 50% to 90% of the landscapes, sporadic permafrost-covering 10% to 50% of the landscapes and isolated permafrost-land covering less than 10% of the landscapes. Discontinuous permafrost constitutes about 7% of the landscapes in the Northern Hemisphere (Brown et al., 1997; Brown et al., 2002; Brown and Haggerty, 1998) (Appendix 1). About 80% of Alaska is occupied by permafrost, including continuous (32%), discontinuous (31%), sporadic (8%), isolated (10%) while other glaciers and ice sheets constitute only 4% of the area (Jorgenson et al., 2008). Thawing and settling of ice-rich terrain (thermokarst) are widely spread in the arctic to subarctic regions where permafrost is found in a discontinuous form (Jorgenson et al., 2001; Halsey et al., 1995). Continuous permafrost has been considered stable because of its much lower annual temperature (i.e. -12 °C to 6 °C) (Nelson et al., 2001).
Permafrost serves as an indicator of possible climate change (Harris et al., 2003). High latitudes experience accouterments of climate change that include degradation of permafrost that has altered the hydrology (Osterkamp et al., 2005; Serreze et al., 2000). Predicted global warming has been focused over the accouterments of increased air temperature and degradation of frozen ice (permafrost) worldwide (Smith, 1983; Anisimov, 1989; Zoltai and Vitt, 1990; Anisimov and Nelson, 1996; Nelson et al., 1993; Houghton et al., 1996). Continued warming will lead to loss of near-surface permafrost in the great aerial extent (Lawrence et al., 2012; Koven et al., 2015; Slater and Lawrence 2013; Schuur et al., 2009; Schuur et al., 2015). Monitoring of permafrost ground is critical as the active layer responds to global warming, relocates snow cover patterns, changes land cover types, and forms thermokarst lakes (Davis, 2001; Ishikawa, 2003).

Permafrost mapping and modeling are important in predicting the change in snow cover areas, land cover patterns and underlying climatic variation (Davis, 2001; Ishikawa, 2003). Permafrost mapping has been initiated in considerable mountainous regions in the world such as Salluit in Canada, the Qinghai-Tibet railway, Yukon’s North Alaska Highway. The permafrost of Alaska is characterized by degradation in Arctic, Alaska and northwest Territories, Canada (Jorgenson et al., 2008; Hegginbottom, 1973; Li et al., 2003; Keller, 1992; Hoelzle et al., 1993; Imhof, 1996; Etzelmüller et al., 2001; Gruber and Hoelzle, 2001; Lugon and Delaloye, 2001; Tannarro et al., 2001; Heginbottom, 2002; Guglielmin et al., 2003).

Permafrost degradation due to climate change and human activities in Alaska has caused major ecological and environmental issues. Permafrost degradation has led to a disruption risk of infrastructures (Smith et al., 2010; Jorgenson et al., 2006), changes in vegetation and hydrology (Walter et al., 2006; Schuur et al., 2009; Jorgenson et al., 2010). Top of the permafrost is ice-rich
(Mackay, 1983; Burn, 1988) and the melting of ground ice degrades permafrost, which may lead to subsidence and accelerate erosion (Mackay, 1970). In central Alaska, ample permafrost is present and is heavily affected by topography and local ecosystems. In addition, the degradation of permafrost is also caused by increased near-surface air temperature and change as winter snow thickness. Anthropogenic activities such as the removal of vegetation for infrastructure construction are most fundamental factors that lead to the degradation of permafrost (Beck et al., 2015; Everdingen, 1998). Also, there are modified models were developed and POP models were validated with permafrost field observation data contain permafrost presence and absence within less than one-meter depth. There are total of 414 field observations that were collect from 1994 to 2012.

The most important factor of permafrost degradation is active layer thickness (ALT). The ALT is the uppermost layer of permafrost that undergoes thawing in summer and freezing in winter. Its thickness fluctuates spatially and temporally (Harris et al., 1988; Muller, 1947; Brown et al., 2000). The ALT plays an important role in permafrost and surficial processes such as ecology and hydrology of permafrost (Liu et al., 2010). The ALT is the rooting zone for plants and acts as a seasonal aquifer for near-surface ground water (Burn, 1998). The depth of ALT greatly varies from a few centimeters to tens of meters, depending on local climatic conditions, thickness of organic layer present, vegetation type, and snow distribution during the winter, and the amount of ground ice (French and Shur, 1977).

1.2 Research Statement

Existing studies have focused and investigated the degradation of permafrost resulting from global climate change and various anthropogenic activities such as land conversion for infrastructure development and industrialization (Houghton et al., 1996; Osterkamp, 1983;
There are two methods for monitoring the dynamics of permafrost at regional scales. The first one is the use of InSAR images to estimate surface deformation-including uplifting and subsidence (Alasset et al., 2010; Zhang et al., 2019; Strozzi et al., 2018; Manandhar, 2019). However, this method is costly since it requires purchasing costly InSAR images in addition to the algorithmic and computational complexity. The second method is the use of optical images (Wang, 2017; Wang 2008; Yaya et al., 2018). This second method is cost-efficient compared to the use of InSAR because the optical images can freely be downloaded. However, there have been only few reports (Dagurov and Chimitdorzhiev, 2005). Wang (2017) first time used optical images to map permafrost presence in Qinghai-Tibet Plateau and found that elevation, SMI, NDVI and albedo were the major driving factors. However, they might not work well in DTA of Alaska because of different environment. Moreover, other optical image derived spectral indices and environmental variables may also significantly contribute the improvement of mapping permafrost presence in DTA. This study will thus concentrate on identifying the main factors or predictors and then mapping the permafrost of DTA using the predictors from optical images with a combination of other auxiliary attributes such as elevation, slope, and aspect derived from a digital elevation model (DEM).
1.3 Objectives and Research Questions

The major objective of this study is to develop a method of mapping permafrost using Landsat images and identify the major factors that affect the mapping accuracy. The research questions are as follow:

Research Question 1: What are the important predictors that significantly improve mapping the presence of permafrost in DTA?

Research Question 2: How does the permafrost vary spatially in the study area?

1.4 Significance of the Study

This work will add on to already available knowledge and skills on permafrost mapping, landscape change, and different factors that account for permafrost accumulation and degradation. It will help understand the long-term change of permafrost in Alaska. Expectedly, the results of this study will act as a baseline for further research and help planners and policy decision-makers to reduce the degradation of permafrost in DTA. In addition, it is also expected that the results can be used for understanding the process and prospects of climate change.
CHAPTER 2
LITERATURE REVIEW

2.1 Permafrost Distribution in Northern High Latitudes and its Degradation

There is a permafrost area of 25.5 million hectares (including glaciers) that occupies about 23% of the land in the Northern Hemisphere. Specifically, permafrost covers about 80% of Alaska including continuous (30%), discontinuous (31%), sporadic (8%), insolated permafrost (10%), and ice sheets (4%) (Jorgenson et al., 2001; Zhang et al., 1999; Brown et al., 1997; Brown and Haggerty, 1998). Permafrost is an indispensable component of many Northern Hemisphere ecosystems that support the land surface, modifies topography, and affects land surface temperature and moisture, subsurface hydrology, rooting zones, and different nutrient cycling (Brown and Grave, 1979; Van Cleve and Viereck, 1983). Permafrost is sensitive to climate change directly through land surface temperature, air temperature, snow cover, moisture, and soil, and indirectly through human turmoil with positive and negative feedbacks (Brown and Grave, 1979; Nelson et al., 2001; Jorgenson et al., 2010).

Permafrost degradation can be influenced by climatic factors in different regions or by surface changes at local and regional scales (Czudek and Demek, 1970; Shur and Osterkamp, 2005). Most studies have mainly focused on the effects of expanded air temperature on degradation of permafrost worldwide (Smith, 1983; Anisimov, 1989; Zoltai and Vitt, 1990; Anisimov and Nelson, 1996; Nelson et al., 1993). The air temperature in the northern latitudes has increased about twice the rate of global mean and is projected to increase further for the 21st century (Arctic Climate Impact Assessment (ACIA, 2004). Climate warming increases ground temperature, thickening of the active layer which refers to a layer or body of unfrozen ground occurring in permafrost regions due to a local anomaly in thermal, hydrological, hydrogeological
or hydrochemical conditions as described by Van Everdingen (1998/2005), and thawing of permafrost (Vaughan et al., 2013). These physical changes affect different infrastructure foundations, hydrology, ecosystems, and overall environment feedbacks (ACIA, 2004). Permafrost-related threats include floods, mass movement, thaw and frost heave. Climate change leads to degradation of permafrost, poses adverse effects on water availability, land subsidence, tourism, and land-related socio-economic consequences (Burger et al., 1999; Nelson et al., 2002; Ford and Smit, 2004) or global scale problems like methane emission from thawing permafrost. Continuous permafrost in the Subarctic has been considered stable due to the much lower mean annual temperature (i.e. -12 °C to 6 °C (Jorgenson et al., 2001; Halsey et al., 1995)).

Permafrost monitoring and mapping using remote sensing techniques in high latitudes and cold mountains have been used for long (Kääb, 2008; Burger et al., 1999; Nelson et al., 2002; Ford and Smit, 2004). Hazards developments and process interactions chain have been ongoing a shift due to atmospheric warning. Kääb et al. (2008) have studied permafrost-related problems and hazards by applying ground-based, airborne and spaceborne remote sensing methods. Digital terrain models (DTMs) were developed from optical stereo, synthetic aperture radar (SAR), and laser scanning were used for investigating permafrost related mass movement (Harris, 2005; Kääb et al., 2005). Accurate and detailed hazards map permafrost with high-resolution space-borne methods have also for long been applied. This study only applied the short-term forecasting of events. This permafrost hazards study requires knowledge about the potential processes (Davies et al., 2006).

The degradation of Arctic zone of continuous permafrost due to abrupt increase of mean annual temperatures. The degradation of ice wedges in northern Alaska indicates the great degradation of permafrost. Jorgenson et al. (2006) evaluated the degradation of ice wedges by
applying field surveys, photo interpretation of a time-series of aerial photography, and spectral characteristics of 1945 and 2001 photography with image-processing software. The study by Jorgenson et al. (2006) revealed the abrupt increase of ice-wedge degradation since 1945, which is beyond the normal rates of change in landscape evolution. This study only quantified recent thermokarst in relatively small areas.

The study was conducted on the Tanana Flats in Central Alaska and revealed that permafrost degradation is widespread and rapid causing large shifts in ecosystems. This abrupt degradation of permafrost has mainly been as a result of increased air temperatures (Smith, 1983; Anisimov, 1989; Zoltai and Vitt, 1990; Anisimov and Nelson, 1996; Nelson et. al., 1993). This study was carried out by collecting field surveys. The area maximized the range of possible permafrost and vegetation conditions and the surface measured along each transect with auto-level rod. Next was sampling points representing the varieties of ecosystems, soil-stratigraphy using the standard methods (Jorgenson et al., 1999). Another method employed was an integrated terrain unit approach for ecosystem classification and mapping that includes separate classification of geomorphology and vegetation. The final method was change detection where permafrost degradation was performed by interpreting the aerial photographs. The radiocarbon dating of the organic material from the base of the peat and tree coring was collected using Beta Analytic, Inc. Wood and peat samples underwent radiocarbon dating. Historical data from mean annual temperature and winter accumulative total snowfall was also collected (ACRC, 1999; D’Arrigo and Jacoby, 1993).

Regional scale permafrost mapping in Mackenzie region in Canada by use of TTOP ground temperature model presented an explicit formulation of climate-permafrost which influence climate, terrain and lithologic factors on the temperature condition and distribution of
permafrost (Smith and Riseborough, 1996; Henry and Smith, 2001). The TTOP model links the ground temperatures to the atmospheric temperature regime with seasonal variations (Lunardini, 1981). The study by Pilon et al. (1991) utilized the TTOP model for regional/sub-regional permafrost modelling and mapping. The TTOP model was calibrated correctly and predicted the occurrence of permafrost at 134 of 154 borehole sites (Pilon et al., 1991). The TTOP model predictions of presence/absence of permafrost was 85%. TTOP model predicted better result for presence/absence of permafrost based on climate and terrain factor and ground temperature; however, TTOP is not fully success due to various factors such as boreholes data, vegetation classes, high topographic influences on local slope and aspect and regional elevation (Pilon et al., 1991).

Permafrost contains about twice as much carbon as the atmosphere (Zhang et al., 1999; Zimov et al., 2006). These carbon quantities stored in frozen soils can be released into the atmosphere anytime due to warming-induced permafrost degradation, which is further enhanced by warming climate (Hodgkins et al., 2014; Hollesen et al., 2015; Schuur et al., 2015). Degradation of permafrost regions shifted from being a sink to a source of CO₂ by the end of the 21st century (Koven et al., 2011). Significant research examining vegetation-derived spatio-temporal response for climate change, permafrost degradation, land surface temperature from satellites images, and land surface models should be conducted to map the permafrost (Kim et al., 2014; Li et al., 2012; Zhang et al., 2003).

Growth of remotely sensed information and their progress from terrestrial vegetation indices can provide extremely useful applications in environmental monitoring, biodiversity conservation, urban green infrastructure, forestry, agriculture, and related fields (Mulla, 2013). Remote sensing-derived vegetation is mainly conducted by obtaining the electromagnetic wave
reflectance information from canopies using different sensors. The reflectance of light spectra from vegetation changes with vegetation type, water content within tissues and other intrinsic factors (Chang et al., 2016). The remotely sensed image-derived vegetation is based on the ultraviolet region which range from 10 to 380 nm and the visible spectra, and is composed of blue (450-495 nm), green (495-570 nm), and red (620-750 nm), and near to mid infrared band (850-1700 nm) (Bin et al., 2016; Gruden et al., 2012). Image-derived vegetation indices as well as other indices and band 10-TIRS I (10.6-11.19 nm) and band 11-TIRS II (11.50-12.51 nm) of Landsat 8 OLI/TIRS were used for calculation of LST.

2.2 Mapping Methods for Permafrost

There is a lot of literature on permafrost mapping that needs to be taken into attention when developing a methodology that proposes to solve the spatial distribution of permafrost in DTA by using remotely sensed data. Yaya et al. (2018) used permafrost influencing factors such as DEM, SMI, NDVI, LST, and albedo generated from remotely sensed data to map the presence and absence of permafrost regions and seasonally frozen ground. They used the decision tree method with other environmental factors to generate a 1 km resolution permafrost map, which was later compared with the maps generated from traditional methods using field surveys. The result showed that the map generated using remotely sensed data had a higher accuracy than those from traditional methods. The mapping accuracy was 85.76% with kappa statistics of 0.71.

In this research, computed regression relationship of permafrost distribution and other influencing factors (elevation, LST, NDVI, SMI, albedo) were analyzed by Wang (2009). The results indicated that the correlation coefficients of four factors (that is elevation, LST, NDVI, and SMI) gave the highest accuracy and indicated that elevation is a governing factor. This method supported the mapping of permafrost by use of certain factors which were good for the
results. The main drawback is that the images used are low resolution and high-resolution images are recommended for better results (Schober et al., 2018.; Wang, 2009). Permafrost distribution study based on GIS and Remote Sensing by Wang (2017) selected model factors (elevation, SMI, NDVI, albedo) that were used in multivariate analysis models and mapped the permafrost presence and absence in Qinghai-Tibet Plateau. This model gives the probabilities of permafrost present in the specified areas. Finally, this method did a comparison of simulation results of permafrost mapping with a published permafrost map, ground temperature inversion model, multivariate analysis model, elevation model, and frost number model had relative error 16.7%, 1.98%, 17.6%, and 10.9% respectively (Wang, 2017; Wang, 2009; Yaya et al., 2018). The main benefits for this study was well-mapped for the distribution of permafrost with the remotely sensed data. The main drawbacks for this study was unable to address more factors. There are more factors also influential factors for the distribution of permafrost. This study had high relative error (Wang, 2009; Wang, 2017).

Other studies that conducted permafrost mapping and its degradation assessment due to climate change were by Jorgenson et al. (2006), Zhang et al. (2014), and Wright et al. (2003). The use of one or more combination of field surveys, satellite images and aerial photointerpretation, spectral characteristics of aerial photography, NEST model (inputs: satellite images, field data, LULC types, ground types), TTOP ground temperature model. All these models have produced better results for permafrost mapping in the high latitudes.

There are variety of methods that have been used to assess uncertainty. Generally, these uncertainties can be tracked as they propagate the methods and develop new methods that reduces the uncertainties. These mentioned methods use remotely sensed images and different vegetation indices are not enough for the research for the distribution of permafrost. To
overcome this drawback addition of more factors and more vegetation indices for the mapping
distribution of permafrost in DTA in this study.

2.3 Remote Sensing Imagery-Derived Vegetation Indices and Factors Affecting
Permafrost Distribution

2.3.1 Normalized Difference Vegetation Index (NDVI)

The NDVI is an indicator to quantify the greenness of plants within certain geographical
areas. It allows enables easy measurement of relative biomass, health of the plants, and indirect
measurement of climatic conditions (Lillesand, 2015). The absorption of chlorophyll in Red band
and reflectance of vegetation in NIR are high (ESRI). Lillesand (2015) stated NDVI as “An
NDVI is often used worldwide to monitor drought, monitor and predict agricultural production,
assist in predicting hazardous fire zones, and map desert encroachment. The NDVI is preferred
for global vegetation monitoring because it helps to compensate for changing illumination
conditions, surface slope, aspect, and other extraneous factors”. Among the several vegetation
indices, the most frequently used vegetation index is the NDVI (Rouse et al., 1973). Mostly
satellite data has been used to calculate NDVI.

2.3.2 Albedo

In the frame of glacier mass balance modeling, albedo is defined as “the ratio of the
reflected flux density to the incident flux density, usually referring either to the entire spectrum of
solar radiation (broadband albedo) or just to the visible part of the spectrum. Spectral albedo is
the albedo at single wavelengths or, more loosely, over a narrow range of wavelengths” (Cogley
et al., 2011). The total radiation plays an important role in biophysical processes that involve the
exchange of energy and mass in the planetary boundary layer and constitutes the main energy
source in the heating of soil, air, and plant evapotranspiration (Cruden et al., 2012; Silva et al.,
Albedo is one of the many parameters used in studying surface energy budget, weather forecast, and many other climatic models such as general circulation model. It is critical for geographical parameter and widely used to study in surface energy budget, weather forecast, and general circulation models (Dickinson, 1983). Santos et al. (2015) stated that the radiation balance has huge relevance in the process of air-soil heating, photosynthesis, and evapotranspiration. Albedo in glaciers determines the amount of energy absorbed by snow/ice surfaces throughout the year. It is important in high latitudes because it balances surface energy. It has been hypothesized that the positive feedback between the changes in surface albedo is associated in snow and ice conditions (i.e., thickness, distribution, and wetness) (Kellog 1975; Shine and Henderson-Sellers 1985; Curry et al. 1995).

### 2.3.3 Land Surface Temperature (LST)

The LST is the temperature of the surface that is measured with direct contact to the measuring instrument. This is also called the skin temperature of the land surface. The LST varies from place to place as well as seasons and different climatic conditions. Mostly, it is found to be varied spatially due to non-homogeneity of land cover and other atmospheric factors (Janak et al., 2012). It has also a great implication in climate. Anthropogenic activities play an important role in changing the land surface temperature. Thus, LST is an essential factor in many areas like global climate change, land use/land cover, geo-biophysical, and input for climate models (Joshi et al., 2012).

Different techniques such as split window technique and single channel technique have been developed to calculate the land surface temperature from satellite images using the
brightness temperature (Janak et al., 2012; Juan et al., 2014; Sun et al., 2010; Juan-Carlos et al., 2008; Rajeshwari & Mani, 2014; Xiaolei et al., 2014; Meijun et al., 2015; Offer et al., 2014).

2.3.4 Soil Moisture Index (SMI)

Soil moisture is the amount of water contained in soil. Soil moisture has an important role in global energy balance, water cycle, and different natural and physiological processes (Mallick et al., 2009; Patel et al., 2009; Vicente et al., 2004). Soil moisture acts as the linkage between hydrological cycle (water) and ecological regime (plants). Interaction between soil-air coupled systems by soil moisture is a most basic phenomenon (Sellers & Schimel, 1993). Land-air coupled system that affects local to regional climate change through energy and water transfers (Skukla et al., 1982; Skukla et al., 1990). Remote sensing technologies widely estimates the soil moisture by using LST and NDVI (Lambin & Ehrlich, 1996).

Several physical and statistical models are used to compute soil moisture. The most well-known physical model for soil moisture is Advanced Integral Equation Model (AIEM) that simulates the radar backscattering coefficient from SAR and various soil parameters (i.e. radar wavelength, polarization, incident angle, soil dielectric constant, and surface roughness) (Chen et al., 2003). Importantly, there are several statistical models that are used to estimates soil moisture developed by Oh et al., and they include the inversion diagram based on the cross-polarized ratio (Y. Oh, 2004; Oh et al., 1992; Oh et al., 2002).

2.3.5 Normalized Difference Snow Index (NDSI)

Snow is the most basic global and meteorological aspect because it plays a significant role in influencing heat regime in local, regional, and global radiation balance. Snow also changes soil characteristics, plant composition, plant community and its structure (Darmody et al., 2004; Löffler 2005). In addition, it is one of the important components of the hydrological
cycle and considered as a major source of fresh water globally (Brown 2000; Yang et al., 2003 & 2010; Zhou and Li 2003; Tong and Velicogna 2010). Snow affects several soil parameters like temperature, moisture, biological activity, soil permeability, and carbon sequestration (Monson et al., 2006; Isard et al., 2007). The NDSI is computed by dividing the difference of reflectance in the Landsat 8 green band (0.53-0.59μm) and the shortwave infrared (1.57-1.65μm). The NDSI threshold of greater than 0.4 is the presence of snow (Hall et al., 1995; Kulkarni et al., 2006; Xiao et al., 2002). Commonly, field surveys data have been used to develop snow maps (Brown and Braaten 1998). Generating snow maps using field data is expensive, time consuming, and tedious. So, remotely sensed datasets are becoming popular during these days, which are open source, freely available, less time consuming, and less expensive (Foody, 2002; Gillanders et al., 2008).

2.3.6 Normalized Difference Water Index (NDWI)

Water is an essential component of ecosystem balance for the sustainability of life on the Earth. Water balances the whole ecosystem and maintains the carbon cycle, climatic variations, etc. It is very important for human and other forms of life that helps to increase the forest and grassland and fluctuation of water amount can affect the land use system (Rover et al., 2012; Alsdorf et al., 2007). Availability of water is conducted by determining the productivity of land, hydropower energy, and irrigation (Wang et al., 2004). The NDWI is the most important and widely used index in measuring how much water is present on the ground. It is a satellite derived index from NIR and SWIR which reflects the water content on land and vegetation canopies, and helps to monitor the drought conditions (Gao, 1996).
2.3.7 Normalized Difference Built-Up Index (NDBI)

Aerial photo and satellite images have become an extensively available source of data for mapping and monitoring land use and land cover change in recent decades. Due to the rapid change of land from forest, agriculture, and bare land turned into urbanization. Data acquired from remote sensing to provide accurate information clues about landscape characteristics and changes into urban areas (Zha et al., 2003). Some of the researchers used Landsat 8 imagery and calculated NDBI by applied Dry-Built-Up Index (DBI) and Dry Bareness Index (DBSI) (Assyakur et al., 2012; He et al., 2010; Bhatti et al., 2014; Faisal et al., 2016). This gives the NDBI in urban areas and bare soil gives low accuracy due to high intensity of homogeneity (Syakur et al., 2012).

2.3.8 Digital Elevation Model (DEM)

The DEM continuously details the elevation and roughness of the Earth surfaces (cf. Evans, 1972; Dikau et al., 1995; 2000) and affects spatial distribution of temperature, vegetation and soil. The thickness of active layer thus varies greatly depending on elevation.
3.1 Study Area

Donnelly Training Area (DTA) is located about 172 kilometers Southeast of Fairbanks, North Star, and west side of Southeast Fairbanks. The total area of DTA is 1,009 Sq. miles. It is bordered by Alaska Range on the south and Fairbanks North Star and Yukon Flats on the north, east, and west (see Figure 3.1 below), which are mostly open land. The southern and western areas of DTA are characterized by steep uplands and alpine environments. This area is bisected by deep streams and wetter lowland environments. The most northern part is characterized by open lowland forests and wetlands (see Figure 3.1).
Figure 3.1: Location map of the study area shown by a Landsat 8 OLI (color composite red, green, and blue: 4, 3, and 2).

According to national land cover database (NLCD, 2016), the major land covered in the study area is dominated by open water, perennial ice/snow, low to high-developed intensity, woody wetlands, some grasslands, shrub, and deciduous to evergreen forest described in Multi-Resolution Land Characteristics (Figure 3.2 below) (https://www.mrlc.gov/data).
The nearest weather station is located in the Fairbanks Alaska, which is northwest of the study area. Average monthly temperature for the years 2014 to 2018 was downloaded from Western Regional Climate Center (https://wrcc.dri.edu/). The maximum recorded monthly temperature from 2014 to 2018 is 64.38 °F while the minimum temperature is 10.64 °F. The monthly average temperature trend from 2014 to 2018 shows normal distribution. The study area depicts the short summertime and has a long cold winter (see Figure 3.3 below).

Figure 3.2: Land Cover in Donnelly Training Area (NLCD, 2016).
The total precipitation was also downloaded from Western Regional Climate Center from 2014 to 2018 showing the maximum total precipitation as 135.85 millimeter on October 2017 with the lowest being 0 millimeter, which occurred in many months in every year from 2014 to 2018. Precipitation has varied greatly from 2014 to 2018 (Figure 3.4).

Figure 3.3: Average Monthly Temperature in Fairbanks, Alaska from 201-2018 (Source: Western Regional Climate Center).

Figure 3.4: Monthly precipitations in Fairbank, Alaska from 2014-2018 (Source: Western Regional Climate Center).
3.2 Datasets

3.2.1 Landsat Images

In this study, Landsat 8 OLI/TIRS Level-1 datasets of paths 67 and 68 with row 15 were downloaded from the USGS Earth Explorer website (https://earthexplorer.usgs.gov/) (Figure 3.5). The temporal range of dataset varies from January 2014 to December 2018 (Table 3.1). The geographic projections of datasets were WGS 1984_UTM_Zone_6N. The data used in this study were cloud cover less than 10% of each row and path. The Landsat 8 OIL/TIRS consists of 11 bands. The spatial resolutions of band 1 through 7 and 9 are 30 meters, bands 10 and 11 are 100 meters, and band 8 is 15 meters. The details of Landsat 8 OIL/TIRS are shown in Table 3.2 below.

Table 3.1: Landsat datasets were used with different periods.

<table>
<thead>
<tr>
<th>Dates</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
<td>X,X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3.2: Summary of band designations and pixel size for Landsat 8 Satellite Sensor.

<table>
<thead>
<tr>
<th>Landsat 8 OLI/TIRS</th>
<th>Wavelengths (Micrometers)</th>
<th>Pixel Size (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band 1 – Ultra Blue (coastal/aerosol)</td>
<td>0.43-0.45</td>
<td>30</td>
</tr>
<tr>
<td>Band 2 – Blue</td>
<td>0.45-0.51</td>
<td>30</td>
</tr>
<tr>
<td>Band 3 – Green</td>
<td>0.53-0.59</td>
<td>30</td>
</tr>
<tr>
<td>Band 4 – Red</td>
<td>0.64-0.67</td>
<td>30</td>
</tr>
<tr>
<td>Band 5 – Near Infrared (NIR)</td>
<td>0.85-0.88</td>
<td>30</td>
</tr>
<tr>
<td>Band 6 – Shortwave Infrared (SWIR) I</td>
<td>1.57-1.65</td>
<td>30</td>
</tr>
<tr>
<td>Band 7 – Shortwave Infrared (SWIR) II</td>
<td>2.11-2.29</td>
<td>30</td>
</tr>
<tr>
<td>Band 8 – Panchromatic</td>
<td>0.50-0.68</td>
<td>15</td>
</tr>
<tr>
<td>Band 9 – Cirrus</td>
<td>1.36-1.38</td>
<td>30</td>
</tr>
<tr>
<td>Band 10 – Thermal Infrared (TIRS) I</td>
<td>10.6-11.19</td>
<td>100 * (30)</td>
</tr>
<tr>
<td>Band 11 – Thermal Infrared (TIRS) II</td>
<td>11.50-12.51</td>
<td>100 * (30)</td>
</tr>
</tbody>
</table>

Note: The table shows each band number and the corresponding wavelength range (in parentheses, micrometers). The exact spectral ranges of each band vary among sensors but are comparable for many applications. This table was adapted from USGS (2017). (Note: TIRS Band 10 and Band 11 are acquired at the 100-meter spatial resolution but are resampled to 30 meters in data product).

3.2.2 Geologic Characteristics of the Donnelly Training Area (DTA)

This Donnelly Training Area is geologically composed of different rock units and soils such as Augen gneiss, coal-bearing sedimentary rocks, some granitic rocks, pelitic quartzite, unconsolidated surficial deposits, and Nenana Gravel (Hults et al., 2015). Predominantly, DTA area is covered by Nenana Gravel. A detailed geological map of the study area is shown in Figure 3.6 (top). Digital Elevation Model (DEM) of the DTA was also downloaded from the
United States Geological Survey (USGS) and elevation ranges from 274.2 to 1877 meters the average level of sea as shown in Figure 3.6 (bottom).

Figure 3.6: Geological rock units in the study area (top) and Elevation map of DTA (bottom).
Slope and aspect maps were generated from DEM. The slope of the Donnelly Training Area ranges from zero to ~58° (Figure 3.7, Table 3.3).

Figure 3.7: Aspect map of the DTA (top) and Slope map of the DTA area (bottom).
Table 3.3: Aspect values from Map.

<table>
<thead>
<tr>
<th>Aspect direction</th>
<th>Value (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>-1.00</td>
</tr>
<tr>
<td>North</td>
<td>0-22.5 and 337.5-360</td>
</tr>
<tr>
<td>Northeast</td>
<td>22.5-67.5</td>
</tr>
<tr>
<td>East</td>
<td>67.5-112.5</td>
</tr>
<tr>
<td>Southeast</td>
<td>112.5-157.5</td>
</tr>
<tr>
<td>South</td>
<td>157.5-202.5</td>
</tr>
<tr>
<td>Southwest</td>
<td>202.5-247.5</td>
</tr>
<tr>
<td>West</td>
<td>247.5-292.5</td>
</tr>
<tr>
<td>Northwest</td>
<td>292.5-337.5</td>
</tr>
</tbody>
</table>

3.2.3 Permafrost Field Observation Data

Permafrost field observation data contain permafrost presence and absence within less than one-meter depth. There are total of 414 field observations that were collect from 1994 to 2012 (Pastick et al., 2014). But, the field data are mainly distributed in the south and southwest parts of the study area. Figure 3.8 shows the locations of the field data for the permafrost presence and absent with a false color composite imager of Landsat 8 OLI bands 6, 5 and 4.
The active layer thickness (ALT) can be defined as the topmost layer of ground that undergoes to annual thawing during the summer and freezing winter in areas underlain by permafrost (van Everdingen, 1998). In DTA, ALT has higher values in the southern, south east, south west as well as river valley and low in norther part because temperature is higher in southern part than in northern part. ALT is highly controlled by surface temperature, physical and thermal properties of the surface cover and subsoil, vegetation, soil moisture, duration, and thickness of seasonal snow cover (Brown et al., 2000). For ALT validation, an ALT map indicating the average ALT values of 2010 to 2019 and developed by Geophysical Institute.
Permafrost Lab (GIPL, 2017) was obtained in 1000 m by 1000 m resolution. The predicted ALT values ranged from 0.47 cm to 31.99cm (Figure 3.9).

Figure 3.9: Predicted Active Layer Thickness of 2010 to 2019 and developed by Geophysical Institute Permafrost Lab (GIPL, 2017).
CHAPTER 4

METHODOLOGY

The methodological flowchart of this study is shown in Figure 4.1 below. The flowchart can be distinctly divided three parts: (1) data acquisition and pre-processing, (2) spectral indices calculation and analysis, (3) model development, and results and decisions.
Figure 4.1: Detail flowchart for the study of permafrost mapping in the Donnelly Training Area.
4.1 Image Pre-processing

Images collected from satellites are affected by different distortions (geometric, radiometric, and atmospheric). These images should be corrected (radiometrically, geometrically, and atmospherically) and obtain better results (Young et al., 2017). The pre-processing of satellite images involves sensor calibration, solar correction, and atmospheric correction (Figure 4.2). Landsat 8 OLI/TIRS level-1 dataset comprises of quantized and calibrated scaled with Digital Numbers (DN) presenting the multispectral data. Landsat 8 satellite datasets were obtained with the Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) available in 16-bit unsigned integer format (ARSC, 2002; Chavez et al., 1996; NASA, 2013). The obtained Landsat 8 level products with cloud cover less than 10% were already geometrically corrected. Thus, in this study the atmospheric correction was conducted.

4.1.1 Atmospheric Correction of Landsat 8 OLI/TIRS

The procedure of retrieving surface reflectance by removing path radiance due to atmosphere from satellite-measured radiance is called atmospheric correction. Thus, atmospheric correction is the process that corrects between atmospheric gases and aerosol, water vapor, ozone conditions, atmospheric scattering, and haze (Pandya, 2015).

Conversion to TOA Radiance (Top of Atmosphere)

Landsat 8 Level-1 data was converted to TOA spectral radiance rescaling factor using the information found in metadata (MTL.txt) file.

\[ L_\lambda = M_L Q_{cal} + A_L \]  
(4.1)

where, \( L_\lambda \) = TOA spectral radiance (Watts/ (m\(^2\) * srad * μm))

\( M_L \) = Band-specific multiplicative rescaling factor from the metadata

(RADIANCE_MULT_BAND_x, where x is the band number)
\( A_L = \text{Band-specific additive rescaling factor from the metadata (RADIANCE_ADD_BAND}_x, \) 
where, \( x \) is the band number

\[ Q_{\text{cal}} = \text{Quantized and calibrated standard product pixel values (DN)} \]


**Conversion to TOA Reflectance**

Landsat 8 Level-1 data reflective band DN's was converted to TOA reflectance rescaling factor using information found in metadata (MTL.txt) file.

\[ \rho_{\lambda'} = M_{\rho} Q_{\text{cal}} + A_{\rho} \]  \hspace{1cm} (4.2)

where,
\[ \rho_{\lambda'} = \text{TOA planetary reflectance, without correction of solar angle. (m}^2 \times \text{srad} \times \mu\text{m)} \]

\( M_{\rho} = \text{Band-specific multiplicative rescaling factor from the metadata} \)

(REFLECTANCE_MULT_BAND_x, where \( x \) is the band number)

\( A_{\rho} = \text{Band-specific additive rescaling factor from the metadata} \)

(REFLECTANCE_ADD_BAND_x, where \( x \) is the band number)

\( Q_{\text{cal}} = \text{Quantized and calibrated standard product pixel values (DN)} \)

TOA reflectance with correction of sun angle is then:

Again, TOA with angle correction and sun angle was used from metadata using the following equation

\[ \rho_{\lambda} = \frac{\rho_{\lambda'}}{\cos (\theta_{SZ})} = \frac{\rho_{\lambda'}}{\sin (\theta_{SE})} \]  \hspace{1cm} (4.3)

where,
\[ \theta_{SE} = \text{Local sun elevation angle. The scene center sun elevation angle in degrees is provided in the metadata (SUN_ELEVATION).} \]

\( \theta_{SZ} = \text{Local solar zenith angle; } \theta_{SZ} = 90^\circ - \theta_{SE} \)
$\rho \lambda'$ = TOA planetary reflectance or final atmospheric correction.

Figure 4.2: Correction processes and units for pre-processing remotely sensed imagery (Young et al., 2017).

4.2 Calculation of Different Optical Remote-sensing Indices from Landsat 8 OLI/TIRS

The following are the remote sensing indices (factors) which will be used to study the permafrost in DTA.

4.2.1 TOA to Brightness Temperature (BT) Conversion

Calculation of brightness temperature (BT) conversion was performed using metadata available while downloading the Landsat 8 OLI/TIRS (NASA, 2008; Artis and Carnahan, 1982).

$$
T = \frac{K_2}{\ln(K_1 L_\lambda + 1)}
$$

(4.4)

where, $T$ = Top of atmosphere brightness temperature (K)

$L_\lambda$ = TOA spectral radiance (Watts/ (m$^2$ * sr * μm))

$K_1$ = Band-specific thermal conversion constant from the metadata (K1_CONSTANT_BAND_x, where, x is band number)
\( K_2 \) = Band-specific thermal conversion constant from the metadata

\( (K2\_CONSTANT\_BAND\_x, \text{ where } x \text{ is band number } (NASA, 2008: \text{Artis and Carnahan, 1982}). \)

### 4.2.2 Normalized Difference Vegetation Index (NDVI)

For the calculation of NDVI, Landsat 8 OLI/TIRS, bands 4 and 5 were used. NDVI enables calculation of relative biomass. The NDVI is used to study drought, agricultural production, and predicting hazards (Lillesand, 2015). Initially, NDVI was calculated for each acquired Landsat 8 data mentioned in table 1. For the calculation of NDVI, the following equation was used:

\[
\text{NDVI} = \frac{(\text{NIR} - \text{RED})}{(\text{NIR} + \text{RED})}
\]

\[
\text{NDVI} = \frac{(\text{TM}_5 - \text{TM}_4)}{(\text{TM}_5 + \text{TM}_4)}
\]

(4.5)

### 4.2.3 Proportion of Vegetation (Pv)

The proportion of vegetation was calculated from the NDVI which was already calculated in equation 4.4. The proportion of vegetation gives the idea about the area that is predominantly vegetation and also indicates the area with less or no vegetation. The proportion of vegetation (Pv) was calculated using the following equation:

\[
\text{P}_v = \left(\frac{\text{NDVI} - \text{NDVI}_{\text{min}}}{\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}}}\right)^2
\]

(4.6)

where NDVI_{min} is the minimum value from NDVI while NDVI_{max} is the maximum value calculated from NDVI.

### 4.2.4 Land Surface Emissivity (ε)

Land surface emissivity (LSE) was derived from the proportion of vegetation utilizing constant numbers as described by Rajeshwari and Mani (2014). Emissivity is denoted by \( \varepsilon \) (Sobrino et al., 2001 & 2005). While LSE is calculated by the following equation:
Emissivity ($\varepsilon$) = 0.004 * $P_v$ + 0.986  \hspace{1cm} (4.7)

Where $P_v$ is the proportion of vegetation.

4.2.5 Land Surface Temperature (LST)

Land surface temperature (LST) is directly related to the surface energy and hydrological balance. The land surface temperature is very important in climate change, hydrological cycle, urban climate, and vegetation monitoring (Chapin et al., 2005; Kalnay et al., 2003; Ramanathan et al., 2001; Wan et al., 2004). In this study, Landsat 8 TIRS thermal bands 10 and 11 were used to calculate the LST. For the calculation of LST, brightness temperature of band 10 and 11, land surface emissivity, and center wavelength of emitted radiance from bands 10 and 11 were followed (Jiménez-Muñoz et al., 2014; Artis and Carnahan, 1982).

$$\text{LST} = \frac{\text{BT}}{1+\left(\frac{\lambda \cdot \text{BT}}{\rho}\right) \ln(\varepsilon)}$$ \hspace{1cm} (4.8)

where $\text{BT} =$ Brightness Temperature

$\varepsilon =$ Land Surface Emissivity

$\lambda =$ wavelength of emitted radiance. For Landsat 8 OLI/TIRS, $\lambda$ for band 10 is 10.8 and $\lambda$ for band 11 is 12.00 (Weng et al., 2004).

$\rho = h \cdot c / k_B (1.438 \times 10^{-2} \text{ m K})$

$h =$ Planck’s constant (6.626 $\times$ 10^{-34} Js)

$c =$ velocity of light (2.998 $\times$ 108 m/s)

$k_B =$ Boltzmann constant (1.38 $\times$ 10^{-23} J/K)

4.2.6 Albedo

Albedo plays an important role in capturing or reflecting the air temperature from the land surface. Increased temperature changes ice or snow cover of the land and has a huge positive feedback in climate (Kellog, 1975). Albedo was calculated using the following equation:
\[ \alpha = \frac{0.356b1 + 0.130b3 + 0.373b4 + 0.005b5 + 0.072b7 - 0.0018}{0.356 + 0.130 + 0.373 + 0.005 + 0.072} \]  

(4.9)

where \( \alpha \) is albedo and 

\( b_1, b_3, b_4, b_5, \) and \( b_7 \) are respective bands of Landsat 8 OLI and TIRS (Liang, 2001).

4.2.7 Soil Moisture Index (SMI)

Soil moisture is a critical component for life, environment, energy-chemical cycles, runoff, evapotranspiration, erosion, and weathering (Wood, 1997; Rozenstein et al., 2014; Hill, Do., 1980; Delworth et al., 1988). SMI can be derived from LST or in combination with LST and NDVI as shown in the following equations:

\[ \text{SMI} = \frac{T_{\text{max}} - T_s}{T_{\text{max}} - T_{\text{min}}} \]  

(4.10)

where, \( T_{\text{max}} \) and \( T_{\text{min}} \) are the maximum and minimum surface temperature for a given soil moisture index data. The SMI is calculated by using the NDVI and LST as follow:

\[ \text{SMI} = \frac{1.1482 \times \text{NDVI} - T_s + 164.74}{-2.2033 \times \text{NDVI} + 211.769} \]  

(4.11)

where NDVI: Normalized Difference Vegetation Index and Ts: Land Surface Temperature (LST).

4.2.8 Normalized Difference Snow Index (NDSI)

Snow is a valuable source of fresh water and the most important component of the hydrological cycle (Brown, 2000; Yang et al., 2003; Zhou and Li, 2003; Tong and Velicogna, 2010). In this study, NDSI was calculated using green and SWIR bands of Landsat 8 OLI/TSIR using following equation:

\[ \text{NDSI} = \frac{\text{Green} - \text{SWIR}}{\text{Green} + \text{SWIR}} \]  

(4.12)

i.e. NDSI = \( \frac{\text{Band3} - \text{Band6}}{\text{Band3} + \text{Band6}} \)
where SWIR (short wave near infrared) or band 6 in Landsat 8 and Green band or band 3 of Landsat 8.

According to Hall et al. (1995), Xiao et al. (2002), and Kulkarni et al. (2006), the threshold value for snow is 0.4. Lower than 0.4 in the region indicates absence of snow and index value greater than 0.4 indicates presence of snow.

**4.2.9 Normalized Difference Built-up Index (NDBI)**

The NDBI is an important index that indicates the intensity of the urbanization. This index can be calculated using SWIR and NIR bands of any multispectral images using the following equation:

\[
\text{NDBI} = \frac{\text{SWIR} - \text{NIR}}{\text{SWIR} + \text{NIR}}
\]

i.e. \( \text{NDBI} = \frac{\text{Band6} - \text{Band5}}{\text{Band6} + \text{Band5}} \)

**4.2.10 Normalized Difference Water Index (NDWI)**

The NDWI is an important index that contains water or moisture in the area. This index is calculated using green and NIR bands of any multispectral images the following below equation:

\[
\text{NDWI} = \frac{\text{GREEN} - \text{NIR}}{\text{GREEN} + \text{NIR}}
\]

i.e. \( \text{NDWI} = \frac{\text{Band3} - \text{Band5}}{\text{Band3} + \text{Band5}} \)

**4.2.11 Simple Ratio (SR)**

The SR is the inverse relationship between the red and near-infrared reflectance of the healthy vegetation. This is the most basic vegetation index (Cohen, 1991; Birth and McVey, 1968). This gives important information about the vegetation biomass or LAI (Haeberli et al., 2002). It is calculated by using the RED and near-infrared bands of Landsat 8 OLI.
SR = $\frac{\text{RED}}{\text{NIR}}$ \hspace{1cm} (4.15)

\[ \text{i.e. SR} = \frac{\text{Band4}}{\text{band5}} \]

4.2.12 Soil Adjusted Vegetation Index (SAVI)

The monitoring of the phenological pattern of the Earth’s vegetated surface tells the length of the growing season and dry-down period (Huete and Liu, 1994). The SAVI is basically incorporated with soil background and atmospheric adjustment factor. The following formula is used to calculate SAVI:

\[ \text{SAVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}} (1 + L) \] \hspace{1cm} (4.16)

where, L is the canopy adjustment factor that depends on the vegetation proportional as well as vegetation density. For bare soil, L is one and dense vegetation is zero. In this case, L is considered as zero.

\[ \text{i.e. SAVI} = \frac{\text{Band5} - \text{Band4}}{\text{Band5} + \text{Band4}} (1 + 0) \] \hspace{1cm} (4.17)

4.2.13 Enhanced Vegetation Index (EVI)

The EVI is very similar to NDVI. This quantifies the vegetation greenness. This corrects canopy noise and is more sensitive to dense vegetation (Jiang et al., 2008). The EVI is calculated as

\[ \text{EVI} = 2.5 \frac{(\text{NIR} - \text{RED})}{(\text{NIR} + 2.4 \text{ RED} + 1)} \] \hspace{1cm} (4.18)

\[ \text{i.e. EVI} = 2.5 \frac{(\text{Band5} - \text{Band4})}{(\text{Band5} + 2.4 \text{ Band4} + 1)} \]
4.2.14 Normalized Burn Ration (NBR)

The NBR is the most widely used vegetation index for mapping of the burn severity (Brewer et al., 2005). It is used for the fire map in fire-affected areas using post fire-image.

\[ NBR = \frac{NIR - SWIR}{NIR + SWIR} \]  \hspace{1cm} (4.19)

\[ \text{i. e. } NBR = \frac{\text{Band5} - \text{Band6}}{\text{Band5} + \text{Band6}} \]

4.2.15 Triangular Vegetation Index (TVI)

The triangular vegetation index was developed by Borge and Leblanc (2000). It describes radiative energy absorbed by pigments as a function of relative difference between red and near-infrared reflectance. The TVI area as triangle is defined in the green peak. It is calculated as follows:

\[ TVI = \frac{1}{2}(120(NIR - GREEN)) - 200(RED - GREEN) \]  \hspace{1cm} (4.20)

\[ \text{i. e. } TVI = \frac{1}{2}(120(\text{Band5} - \text{Band3})) - 200(\text{Band4} - \text{Band3}) \]

4.2.16 Reduced Simple Ratio (RSR)

Chen et al. (2002) elevated the simple ratio algorithm which include information from short-wavelength infrared (SWIR) band that can be found in vegetation sensor. The sensor map was used to estimate the distribution of LAI in Canada. The RSR is calculated as

\[ RSR = \frac{NIR}{RED} \left( \frac{SWIR - SWIR\text{min}}{SWIR\text{max} + SWIR\text{min}} \right) \]  \hspace{1cm} (4.21)

where, SWIR\text{max} is the short-wavelength infrared maximum and SWIR\text{min} is short-wavelength infrared minimum.

\[ \text{i. e. } RSR = \frac{\text{Band5} - \text{Band6}\text{min}}{\text{Band4} \left( \text{Band6max} + \text{Band6min} \right)} \]
The RSR is so important because the land cover mapping is more accurate over the leaf-area index in mixed covered areas.

4.2.17 Visible Atmospherically Resistant Index (VARI)

The atmospherically resistant vegetation index of visible (that is green) is very sensitive to atmospheric effects. This is an estimation of vegetation fraction with an error of <10% (Rundquist, 2002). The VARI is calculated as

\[
\text{VARI} = \frac{\text{GREEN} - \text{RED}}{\text{GREEN} + \text{RED} - \text{Blue}}
\]

i. e. VARI = \frac{\text{Band3} - \text{Band4}}{\text{Band3} + \text{Band4} - \text{Band2}}

4.2.18 Active Layer Thickness (ALT)

The ALT is the thickness of the surface soil layer which thaws during summer and freezes in the winter. The ALT is an important indicator for monitoring permafrost degradation in the Arctic region. Increase or decrease of permafrost depends on the ALT (Hinzman et al., 2005; White et al., 2007). Degradation of permafrost could have serious effects on biological, biogeochemical, hydrologic, and landscapes changes (Hinzman et al., 2005; White et al., 2007). Assuming the surface subsidence is caused purely by the phase change of pore ice water in an active layer, the change in surface subsidence \((d\delta)\), and due to underlying permafrost thawing (Figure 4.3).
Surface subsidence is caused purely by the phase change of pore ice water in an active layer. The surface subsidence ($d\delta$), caused by underlying permafrost thawing can be described as follows:

$$d\delta = PS \left( \frac{\rho_w - \rho_i}{\rho_i} \right) dh$$

(4.23a)

where $P$ is the soil porosity, $S$ is soil moisture fraction of saturation, $\rho_w$ is density of water (in kilograms per cubic meter) that is 997 kg/m$^3$, $\rho_i$ is density of ice (in kilograms per cubic meter) i.e. 934 kg/m$^3$, and $dh$ is incremental thickness of the thawed soil column (in meters). Integrating both sides of equation gives:

$$\delta = \int_0^H d\delta = \int_0^H PS \left( \frac{\rho_w - \rho_i}{\rho_i} \right) dh$$

(4.23b)

where, $H$ is active layer thickness, and $\delta$ is overall subsidence in thawing season. $d\delta$ is incremental change in surface subsidence (in meters). This is mapped by using DInSAR technique, which is based on Sentinel-1 SAR pairs (Lin et al., 2012). The $S$ is soil moisture.
saturation. This is estimated based on the soil moisture index calculated from Landsat 8 imagery. In this study, the active layer thickness of the optical remote sensing images was calculated as stated by Lin et al. (2012):

\[
H = \frac{\rho_i}{PS(\rho_w - \rho_i)} * \delta
\]  

(4.23c)

4.3 Correlation Analysis and Selection of Model Factors

Many factors affect the permafrost distribution, including weather, topographic feature, soil properties, vegetation, land use and land cover, etc. In the previous studies, elevation, soil moisture index (SMI), albedo, and NDVI were taken as model predictors for determining the probability of permafrost presence based on equation 4.24 introduced by Wang (2017). In addition to the used factors, however, other variables such as LST, NDWI, NDSI, NDBI, SR, PVI, SAVI, NBR, EVI, TVI, and slope, aspect, active layer thickness, may be significantly correlated with the probability of permafrost presence. In this study, all the predictors were used and selected to improve the model for predicting the probability of permafrost presence (POP) based on their correlations with the presence of permafrost and prediction accuracy. The multicollinearity among the model predictors was also analyzed using Variance Inflation Factor (VIF) (Belsley et al., 2005). The modified model was different from the original model proposed by Wang (2017). The optical remote sensing variables were utilized in these models by taking their average values from January 2014 to December 2018 of Landsat 8 OLI/TIRS data, instead of using the values from one single time image. The values of remote sensing variables from individual time images may vary over space and time. Their average values of the model predictors from multi-temporal images may lead to stable predictions. In addition, a logistic model means the nonlinear relationships of the dependent variable with the predictors. However, a logistic equation can be converted to a linear model by log transformation. That is,
\[
\text{POP} = \frac{1}{(1 + \exp(a_0 + a_1 x_1 + a_2 x_2 + \cdots a_m x_m))}
\]

Converted to
\[
\ln\left(\frac{1 - \text{POP}}{\text{POP}}\right) = a_0 + a_1 x_1 + a_2 x_2 + \cdots a_m x_m
\]

Thus, the correlation coefficients of \(\ln((1-\text{POP})/\text{POP})\) with the predictors were calculated and their significant differences from zero were statistically tested at the significant level 0.05.

4.4 Permafrost Distribution Mapping of DTA

The probability model (Equation 4.24) described by Wang (2017) was first used in mapping the permafrost distribution of DTA:

\[
\text{POP} = \frac{1}{(1 + \exp(0.001x_1 + 12.38x_2 - 1.34x_3 + 0.55x_4 - 9.73))} \quad (4.24)
\]

where, \(\text{POP}\) is a probability of permafrost and \(x_1, x_2, x_3,\) and \(x_4\) respectively represent the elevation, SMI, NDVI, and albedo. This equation was fully valid and led to accurate predictions of permafrost presence in Qinghai-Tibet Plateau. But it might not work in DTA of Alaska because of different environment.

In this study, the logistic model was first used to develop the model that accounts for POP. Logistic stepwise regression with the VIF equal to 10 was utilized to select the driving factors. Because ALT significantly affects the POP, similarly, the model that explains the relationship of ALT with the driving factors was also derived using linear stepwise regression with the VIF equal to 10.

4.5 Validations of Models

Based on 414 field observations, the accuracy of predictions from each of the models obtained was assessed using the adjusted coefficient of determination and relative root mean
square error (relative RMSE) between the estimated and observed values, and Akaike’s information criterion AIC (Akaike, 1973, 1974). The AIC is a one kind of model information selection which analyzes the true calculated model with predicted model. In this study, AIC can be estimated using the following equation.

\[
\text{AIC} = N \ln \left( \frac{\text{ESS}}{N} \right) + 2K
\]

(4.25)

where, \( N \) is the number of random observations, \( \text{ESS} \) is the Residual Sum of Squares, and \( K \) is the number of model parameters. The adjusted coefficient of determination, relative RMSE and AIC were also used to compare the performance of the predictions from the models.
CHAPTER 5

RESULTS

5.1 Optical Remote Sensing Factors

Different optical remote sensing factors and vegetation indices were developed for modelling POP and analyzing how the factors improve the predictions of POP in DTA of Alaska.

5.1.1 Vegetation, Soil and Water Relevant Indices

5.1.1.1 Vegetation Relevant Indices

Normalized Difference Vegetation Index: The NDVI gives a measure of the amount and vigor of vegetation in a study area. The Average NDVI values of DTA based on the images from 2014 to 2018 ranges from -0.61 to 0.63 (Figure 5.1). According to Sturm et al. (2001), higher values of NDVI indicates greater vigor and amounts of vegetation. In general, NDVI values for vegetated lands range from 0.1 to 0.7. The values greater than 0.5 imply densely vegetated areas, while the values between 0.2 to 0.1 imply snow, inland water body, desert, and exposed soils (Sturm et al., 2001).
Figure 5.1: The average normalized difference vegetation index of DTA.

Enhanced Vegetation Index: The EVI aims to improve the NDVI by optimizing the vegetation signals in higher NDVI areas. This blue region of the spectrum helps to compensate for soil background and reduces atmospheric influenced (that is aerosol scattering) (Huete, Justice, & Van Leeuwen, 1999). The EVI value was ranged from -1 to 1. The calculated EVI in DTA was found to be -0.16 to 0.72 (Figure 5.2).
Visible Atmospherically Resistant Index: This is also an enhancement of NDVI to make it relatively more resistant to atmospheric factors (that is aerosol levels). The VARI uses reflectance values from blue region and correct red reflectance for atmospheric scattering effects (Kaufman & Tanre, 1992). It is most useful in the regions with high atmospheric aerosol content. The value was found ranging from -0.71 to 0.37 in the DTA.

Figure 5.2: Enhanced Vegetation Index (EVI) in the Donnelly Training Area, Alaska.
Simple Ratio and Reduced Simple Ratio: The SR values in DTA range from -0.47 to 0.542 (Figure 5.4). Chen et al. (2002) modified the SR algorithm that includes more information from short-wavelength infrared for the vegetation. This also used the sensor to map the geographical distribution of leaf-area index in Canada. The RSR in the Donnelly Training Area ranges from -0.63 to 0.64 (Figure 5.5).

Figure 5.3: Visible Atmospherically Resistant Index (VARI) in the Donnelly Training Area.
Figure 5.4: Simple Ratio (SR) in the Donnelly Training Area, Alaska.

Figure 5.5: Reduced Simple Ration (RSR) in the Donnelly Training Area, Alaska.
Proportion of Vegetation (Pv): The Pv values range almost 0 to 0.99 (Figure 5.6). The higher values of Pv indicate dense vegetation cover and lower values indicate sparse vegetation cover. In this study, lower Pv values were observed in southern parts of DTA whereas higher values were observed in the northern parts. The proportion of vegetation was also dominant in the Southeast side of the DTA. The Pv values are intermediate results used for calculation of LST.

![Pv Map of Donnelly Training Area, Alaska](image)

Figure 5.6: Proportion of vegetation of Donnelly Training Area in 02/10/2018.

Triangular Vegetation Index: The values of TVI calculated in DTA were from -1017.21 to 618.74 (Figure 5.7). This is a good indication of the chlorophyll estimation and its total reflectance.
5.1.1.2 Soil and Water Relevant Indices

Soil Moisture Index: The SMI gives the function as the linkage between water, ecological regime, and atmosphere (Vicente-Serrano et al., 2004; Engman Chauhan, 1995; Sellers et al., 1993). The average soil moisture values based on the images from January 2014 to December 2018 were 0.75 to 0.83 (Figure 5.8). As defined by Goward et al. (2002), SMI below values 0.34 indicates that the lands have mild to severe desertification while SMI values higher than 0.34 indicates that the lands are wet to very wet. Thus, the results in the DTA indicate that DTA was wet to very wet.
Soil Adjusted Vegetation Index: This index adjusts for the influence of variation in soil background of Red and NIR relationship for this study area with the help of vegetation cover (Huete, 1998). The SAVI value was ranged from -1.01 to 0.82 in this study area (Figure 5.9).
Figure 5.9: The average Soil Adjusted Vegetation Index (SAVI) in the Donnelly Training Area, Alaska.

**Normalized Difference Water Index:** The NDWI shows the water content and vegetation canopies. In DTA, the average NDWI values range from -0.46 to 0.37 based on the images from January 2014 to December 2018 and rivers and snow-covered areas have high NDWI (Figure 5.10).
Figure 5.10: The Average Normalized Difference Water Index (NDWI) in the Donnelly Training Area.

Normalized Difference Snow Index: The NDSI is regarded as the most important and significant role in heat regimes, local and global radiation balance (Darmody et al., 2004; Löffler 2005). The average NDSI values in DTA vary from -0.26 to 0.54 based on the images from January 2014 to December 2018 (Figure 5.11). As recommended by Hall et al. (1995), Kulkarni et al. (2006), and Xiao et al. (2002), the optimal values of NDSI greater than 0.40 highlights the presence of snow and the values lower than 0.40 indicates the presence of water and soil. The results in DTA indicates that the NDSI value ranges from 0.4 to 0.54 represents the presence of snow and while lower than 0.4 values indicate land and water body.
Figure 5.11: The Average Normalized Difference Snow Index in the Donnelly Training Area.

5.1.2 Normalized Burn Ratio: The values of NBR change from -0.25 to 0.41 in DTA (Figure 5.12). This is used for mapping the burn severity and the fire-affected areas using the single post-fire values. It combines the reflectance in the near infrared and the short wavelength of infrared bands. Based on following information (table 5.1) developed by Boer et al. (2008), DTA is considered to be moderately low severity burning (Cocke et al., 2005).
Table 5.1: Burning classes based on Normalized Burn Ratio (NBR) (Boer et al., 2008)

<table>
<thead>
<tr>
<th>NBR</th>
<th>Burn Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;-0.25</td>
<td>High post-fire regrowth</td>
</tr>
<tr>
<td>-0.25 to -0.1</td>
<td>Low post-fire regrowth</td>
</tr>
<tr>
<td>-0.1 to +0.1</td>
<td>Unburned</td>
</tr>
<tr>
<td>0.1 to 0.27</td>
<td>Low severity burn</td>
</tr>
<tr>
<td>0.27 to 0.44</td>
<td>Moderate-low severity burn</td>
</tr>
<tr>
<td>0.44 to 0.66</td>
<td>Moderate-high severity burn</td>
</tr>
<tr>
<td>&gt;0.66</td>
<td>High severity burn</td>
</tr>
</tbody>
</table>

Figure 5.12: The Average Normalized Burn Ratio (NBR) in the Donnelly Training Area, Alaska.
5.1.3 Land Surface Emissivity and Land Surface Temperature

Land Surface Emissivity: In the DTA, LSE values range from 0.986 to 0.999 (Figure 5.13). The LSE was very high in the Northeastern to Northern and Northwest part whereas low LSE values were observed in the water bodies, rivers, and lakes. The LSE is also an intermediate result for the calculation of LST.

![Emissivity Map of Donnelly Training Area, Alaska](image)

*Figure 5.13: The Average Land Surface Emissivity in the Donnelly Training Area.*

Land Surface Temperature: The LST changes with climatic conditions and different human activities. It directly depends on land covers types such as urban areas, forested lands, bare lands, weather conditions, and different terraces. In DTA, the average LST values based on the images from January 2014 to December 2018 were calculated ranged from -11.96 °C to 6.29 °C (Figure 5.14). The average annual temperature described by Douglas et al. (2016) in the interior of Alaska was be -3.3 °C, while the mean summer temperature was 20 °C, and the mean winter...
temperature was -20 °C.

Figure 5.14: The Average Land Surface Temperature (LST) in the Donnelly Training Area.

5.1.3 Other Surface Related Indices

**Albedo:** Albedo is a controlling factor of the earth’s surface energy budget and it is responsible for the reflection or absorption of the solar radiation from the earth’s surface (Liang et al., 2010). The average albedo values based on the images from January 2014 to December 2018 were 0.05 to 0.59 (Figure 5.15). Low albedo mostly indicates water, dry soil, sandy soil, and agricultural crops. Medium to high albedo indicates the area with melting to fresh snow. The study in the North Slope Alaska in 1992 was conducted with the help of Ground-based and Satellite-Derived measurements for surface albedo values that are divided into five seasonal and temporal periods (Maykut and Church, 1973). First, winter stationary period starts from mid-October to May and
the ground surface is completely covered by dry snow. In this period, albedo varies from 0.7 to 0.9, and sometimes drops below 0.7 (Maykut and Church, 1973). The second is the spring snowmelt period, which starts from late May to early June. Snow starts melting and surface albedo is decreases from 0.8 to 0.6 or below. The third season is post-snowmelt period; this is the situation of after disappearing of snow from the ground surface. The surface is generally wet, and the albedo is at its lowest in this period. The albedo ranges from 0.5 to 0.1. This time, the surface is covered with standing water (Maykut and Church, 1973; Weller and Holmgren 1974). The fourth period is the summer stationary period and it lasts from the end of post-snowmelt to the onset of freeze-up. The surface represents drier periods and albedo starts to increase. The albedo of 0.10 is wet tundra while 0.20 is dry tundra. The fifth season is the autumn freeze-up period which starts from late August to mid-October (Maykut and Church, 1973). In this period, stable snow is fully established. The albedo in the North Slope Alaska greater than 0.90 indicates dry snow and can be as low as 0.05 for tundra. The results in DTA represent the tundra to high snow time (Maykut and Church, 1973).
Figure 5.15: The Average Albedo in the Donnelly Training Area.

**Normalized Difference Built-up Index:** The NDBI is much less in the northeast sides of DTA, where the Army training buildings and other infrastructures are located. This result show that higher NDBI values exist mostly in higher elevation and river areas. This might be due to the mixing of barren land, river deposited sand, and gravels which give false result. The NDBI values in DTA ranges from -0.53 to 0.21 (Figure 5.16).
Figure 5.16: The Average Normalized Difference Built-up Index (NDBI) in the Donnell Training Area.
5.1.4 Active layer thickness and probability of permafrost

**Active Layer Thickness:** The ALT is the top layer of the ground, which has the annual thawing, and freezing area underlain by permafrost (Romanovsky & Osterkamp, 1997; Bonnaventure & Lamoureux, 2013; Harris et al., 1988). This ALT plays an important role in land surface processes in the cold areas. The ALT in the DTA ranges from 1.64 cm to 38.88 cm (Figure 5.17). The high values indicate the area that goes seasonally thawing above permafrost and present in southeast to southern part (Harris et al., 1988), whereas low values present in northeast to northern part as well as western part in DTA.

*Figure 5.17: The Active Layer Thickness (ALT) in the Donnelly Training Area, Alaska.*
Permafrost Probability Map in DTA: The POP map of DTA was calculated based on the model factors described by Wang (2017) model. The model factors are elevation, SMI, NDVI, and albedo. The obtained permafrost probabilities ranged from 7.85% to 55.09% (Figure 5.18).

Figure 5.18: The distribution probabilities of permafrost presence (POP) in Donnelly Training Area, Alaska, based on the model proposed by Wang (2017).

5.2 Contribution of Factors

In this study, a total of the 500 points were randomly selected to extract the mentioned optical remote sensing derived factors and conduct correlation and regression analyses (Figure 5.19).
Figure 5.19: The randomly selected 500 points in the DTA area, Alaska.

5.2.1 Correlation Analysis

The correlation analyses of Ln((1-POP)/POP) and ALT with the driving factors, and among the other factors were performed using SPSS 19.0 (Statistical Package for the Social Sciences). The significance of the Pearson correlations was assessed at the significant level of 0.05. Table 5.2 and Figure 5.20 give the Pearson’s Correlation coefficients.
Table 5.2: Correlation summary of Ln((1-POP/POP)) and ALT with the remote sensing image derived factors and among the factors

(Note: * correlation coefficients that are statistically significantly different from zero, and ELE. is Elevation)

<table>
<thead>
<tr>
<th></th>
<th>ELE.</th>
<th>SMI</th>
<th>NDVI</th>
<th>ALBEDO</th>
<th>LST</th>
<th>NDWI</th>
<th>NDSI</th>
<th>NDBI</th>
<th>SLOPE</th>
<th>ASPECT</th>
<th>SR</th>
<th>SAVI</th>
<th>EVI</th>
<th>NBR</th>
<th>TVI</th>
<th>RSR</th>
<th>VARI</th>
<th>ALT</th>
<th>POP</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELE.</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMI</td>
<td>0.03</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>-0.08</td>
<td>-0.07</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALBEDO</td>
<td>0.53*</td>
<td>0.14*</td>
<td>-0.24</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LST</td>
<td>-0.06</td>
<td>-0.99*</td>
<td>0.21*</td>
<td>-0.18*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDWI</td>
<td>0.02</td>
<td>0.34*</td>
<td>-0.74*</td>
<td>-0.09</td>
<td>-0.48*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDSI</td>
<td>0.19*</td>
<td>0.56*</td>
<td>-0.51*</td>
<td>0.20*</td>
<td>-0.62*</td>
<td>0.68*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NDBI</td>
<td>0.09</td>
<td>-0.24*</td>
<td>-0.43*</td>
<td>0.28*</td>
<td>0.16*</td>
<td>0.16*</td>
<td>-0.32*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLOPE</td>
<td>0.53*</td>
<td>-0.15*</td>
<td>0.07</td>
<td>0.01</td>
<td>0.13*</td>
<td>0.11*</td>
<td>0.01</td>
<td>0.003</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ASPECT</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
<td>-0.04</td>
<td>-0.01</td>
<td>0.04</td>
<td>-0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>-0.07</td>
<td>0.04</td>
<td>-0.70*</td>
<td>-0.03</td>
<td>-0.15*</td>
<td>0.66*</td>
<td>0.42*</td>
<td>0.30*</td>
<td>0.05</td>
<td>0.04</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SAVI</td>
<td>0.12*</td>
<td>-0.01</td>
<td>0.78*</td>
<td>0.11*</td>
<td>0.14*</td>
<td>-0.75*</td>
<td>-0.35*</td>
<td>-0.47*</td>
<td>-0.05</td>
<td>-0.05</td>
<td>-0.93*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EVI</td>
<td>0.29*</td>
<td>-0.05</td>
<td>0.59*</td>
<td>0.45*</td>
<td>0.15*</td>
<td>-0.75*</td>
<td>-0.29*</td>
<td>-0.34*</td>
<td>0.02</td>
<td>-0.06</td>
<td>-0.72*</td>
<td>0.87*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NBR</td>
<td>-0.12*</td>
<td>0.26*</td>
<td>0.25*</td>
<td>-0.41*</td>
<td>-0.21*</td>
<td>0.10</td>
<td>0.47*</td>
<td>-0.89*</td>
<td>0.09</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.14*</td>
<td>0.01</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TVI</td>
<td>-0.35*</td>
<td>0.05</td>
<td>0.23*</td>
<td>-0.69*</td>
<td>0.01</td>
<td>0.16*</td>
<td>0.22*</td>
<td>-0.62*</td>
<td>0.04</td>
<td>-0.002</td>
<td>0.09</td>
<td>-0.01</td>
<td>-0.29*</td>
<td>0.75*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSR</td>
<td>-0.03</td>
<td>0.12*</td>
<td>0.72*</td>
<td>-0.16*</td>
<td>0.02</td>
<td>-0.51*</td>
<td>-0.03</td>
<td>-0.84*</td>
<td>-0.02</td>
<td>-0.05</td>
<td>-0.67*</td>
<td>0.83*</td>
<td>0.64*</td>
<td>0.60*</td>
<td>0.46*</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VARI</td>
<td>-0.17*</td>
<td>-0.01</td>
<td>-0.76*</td>
<td>-0.21*</td>
<td>-0.12*</td>
<td>0.76*</td>
<td>0.31*</td>
<td>0.47*</td>
<td>0.02</td>
<td>0.05</td>
<td>0.87*</td>
<td>-0.99*</td>
<td>-0.93*</td>
<td>-0.15*</td>
<td>0.06</td>
<td>-0.82*</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALT</td>
<td>0.28*</td>
<td>0.26*</td>
<td>-0.05</td>
<td>0.36</td>
<td>-0.27*</td>
<td>-0.02</td>
<td>0.08</td>
<td>0.17*</td>
<td>0.17*</td>
<td>0.02</td>
<td>-0.12*</td>
<td>0.12*</td>
<td>0.17*</td>
<td>-0.21*</td>
<td>-0.31*</td>
<td>-0.04</td>
<td>-0.14*</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>POP</td>
<td>-0.81*</td>
<td>-0.24*</td>
<td>0.54*</td>
<td>-0.55*</td>
<td>0.34*</td>
<td>-0.46*</td>
<td>-0.28*</td>
<td>-0.38*</td>
<td>-0.03</td>
<td>-0.27*</td>
<td>0.29*</td>
<td>0.08</td>
<td>0.24*</td>
<td>0.45*</td>
<td>0.39*</td>
<td>-0.24*</td>
<td>0.58*</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>
(a) Scatter plot of Elevation and presence of permafrost

(b) Scatter plot of SMI and presence of permafrost

(c) Scatter plot of NDVI and presence of permafrost

(d) Scatter plot of Albedo and presence of permafrost
(e) Scatter plot of LST and presence of permafrost

(f) Scatter plot of NDWI and presence of permafrost

(g) Scatter plot of NDSI and presence of permafrost

(h) Scatter plot of NDBI and presence of permafrost
(i) Scatter plot of Slope and presence of permafrost

(j) Scatter plot of Aspect and presence of permafrost

(k) Scatter plot of SR and presence of permafrost

(l) Scatter plot of SAVI and presence of permafrost
(m) Scatter plot of EVI and presence of permafrost

(n) Scatter plot of NBR and presence of permafrost

(o) Scatter plot of TVI and presence of permafrost

(p) Scatter plot of RSR and presence of permafrost
(q) Scatter plot of VARI and presence of permafrost

(r) Scatter plot of ALT and presence of permafrost

Figure 5.20: Scatter plot with fit line of Presence of permafrost (POP) with predictor.
Table 5.3: Interpreting the size of a Correlation Coefficient (Mukaka, 2012).

<table>
<thead>
<tr>
<th>Size of Correlation</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90 to 1.00 (-0.90 to-1.00)</td>
<td>Very high positive (negative) correlation</td>
</tr>
<tr>
<td>0.70 to 0.90 (-0.70 to-0.90)</td>
<td>High positive (negative) correlation</td>
</tr>
<tr>
<td>0.50 to 0.70 (-0.50 to-0.70)</td>
<td>Moderate positive (negative) correlation</td>
</tr>
<tr>
<td>0.30 to 0.50 (-0.30 to-0.50)</td>
<td>Low positive (negative) correlation</td>
</tr>
<tr>
<td>0.00 to 0.30 (0.00 to-0.30)</td>
<td>Negligible correlation</td>
</tr>
</tbody>
</table>

The scatter plots of \( \ln((1-\text{POP})/\text{POP}) \) and predictive factors were presented in Figure 5.20. Permafrost is moderately to highly correlated with elevation, NDVI, albedo and ALT (Table 2 and Table 5.3). Overall, \( \ln((1-\text{POP})/\text{POP}) \) has relatively low correlation with LST, NDWI, NDSI, slope, TVI and RSR, and other factors have negligible contributions to the presence of permafrost in DTA.

Pearson’s correlation coefficients between \( \ln((1-\text{POP})/\text{POP}) \) and most of the factors, including elevation, NDVI, albedo, ALT, LST, NDWI, NDSI, slope, TVI, RSR, SMI, NDBI, SR, SAVI, NBR and VARI, were statistically significantly different from zero at the significant level of 0.05, indicating potentially significant contributions to prediction of POP. Also, Spearman’s Rank correlation coefficients were calculated, there are not much differences with Pearson correlations. The elevation has the highest negative correlation (i.e. 0.81) with POP (Table 5.2 and Figure 5.20), then ALT, albedo, NDVI, NDWI, NDSI. TVI, RSR, slope, and LST. On the other hand, \( \ln((1-\text{POP})/\text{POP}) \) has no significant correlation with aspect and EVI, implying that these two variables have statistically negligible effects to improving the prediction of POP.
5.3. Mapping probability of permafrost presence

Table 5.4 shows the adjusted R Square and F values and the selected driving factors of the obtained six models. The coefficients of the models with VIF values are listed in Table 5.5.

Table 5.4: Summary of models obtained for prediction of probability of permafrost presence using stepwise regression.

<table>
<thead>
<tr>
<th>Model</th>
<th>R²</th>
<th>Adjusted R²</th>
<th>Sig. F</th>
<th>Independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.840</td>
<td>0.839</td>
<td>0.000</td>
<td>ELEVATION, NDVI, LST, TVI, ALT</td>
</tr>
<tr>
<td>2</td>
<td>0.845</td>
<td>0.843</td>
<td>0.000</td>
<td>ELEVATION, NDVI, LST, TVI, ALT, SLOPE</td>
</tr>
<tr>
<td>3</td>
<td>0.853</td>
<td>0.851</td>
<td>0.003</td>
<td>SLOPE, RSR</td>
</tr>
<tr>
<td>4</td>
<td>0.858</td>
<td>0.856</td>
<td>0.000</td>
<td>SLOPE, RSR, SMI</td>
</tr>
<tr>
<td>5</td>
<td>0.862</td>
<td>0.859</td>
<td>0.001</td>
<td>SLOPE, RSR, SMI, NBR</td>
</tr>
<tr>
<td>6</td>
<td>0.865</td>
<td>0.862</td>
<td>0.002</td>
<td>SLOPE, RSR, SMI, NBR, NDSI</td>
</tr>
</tbody>
</table>
Table 5.5: The coefficients and variation inflation factors (VIF) of the models obtained for prediction of probability of permafrost presence using stepwise regression.

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables</th>
<th>Coefficients</th>
<th>Variation Inflation Factor (VIF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ELEVATION</td>
<td>0.001</td>
<td>2.058</td>
</tr>
<tr>
<td></td>
<td>NDVI</td>
<td>-1.189</td>
<td>1.109</td>
</tr>
<tr>
<td></td>
<td>LST</td>
<td>-0.061</td>
<td>1.062</td>
</tr>
<tr>
<td></td>
<td>TVI</td>
<td>-0.001</td>
<td>1.201</td>
</tr>
<tr>
<td></td>
<td>ALT</td>
<td>-0.012</td>
<td>1.908</td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>0.087</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ELEVATION</td>
<td>0.001</td>
<td>2.841</td>
</tr>
<tr>
<td></td>
<td>NDVI</td>
<td>-1.197</td>
<td>1.110</td>
</tr>
<tr>
<td></td>
<td>LST</td>
<td>-0.065</td>
<td>1.100</td>
</tr>
<tr>
<td></td>
<td>TVI</td>
<td>-0.001</td>
<td>1.302</td>
</tr>
<tr>
<td></td>
<td>ALT</td>
<td>-0.013</td>
<td>1.927</td>
</tr>
<tr>
<td></td>
<td>SLOPE</td>
<td>0.005</td>
<td>1.586</td>
</tr>
<tr>
<td>3</td>
<td>Constant</td>
<td>0.098</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ELEVATION</td>
<td>0.001</td>
<td>3.117</td>
</tr>
<tr>
<td></td>
<td>NDVI</td>
<td>-1.544</td>
<td>2.579</td>
</tr>
<tr>
<td></td>
<td>LST</td>
<td>-0.061</td>
<td>1.132</td>
</tr>
<tr>
<td></td>
<td>TVI</td>
<td>-0.001</td>
<td>1.819</td>
</tr>
<tr>
<td></td>
<td>ALT</td>
<td>-0.014</td>
<td>1.930</td>
</tr>
<tr>
<td></td>
<td>SLOPE</td>
<td>0.007</td>
<td>1.735</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>RSR</td>
<td>0.107</td>
<td>3.110</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Constant</td>
<td>51.012</td>
<td></td>
</tr>
<tr>
<td>ELEVATION</td>
<td>0.001</td>
<td>3.145</td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>-1.068</td>
<td>6.322</td>
<td></td>
</tr>
<tr>
<td>LST</td>
<td>-0.354</td>
<td>2.569</td>
<td></td>
</tr>
<tr>
<td>TVI</td>
<td>-0.001</td>
<td>1.821</td>
<td></td>
</tr>
<tr>
<td>ALT</td>
<td>-0.014</td>
<td>1.933</td>
<td></td>
</tr>
<tr>
<td>SLOPE</td>
<td>0.005</td>
<td>1.911</td>
<td></td>
</tr>
<tr>
<td>RSR</td>
<td>0.139</td>
<td>3.480</td>
<td></td>
</tr>
<tr>
<td>SMI</td>
<td>-65.378</td>
<td>2.510</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Constant</td>
<td>47.103</td>
<td></td>
</tr>
<tr>
<td>ELEVATION</td>
<td>0.001</td>
<td>3.167</td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>-1.175</td>
<td>6.633</td>
<td></td>
</tr>
<tr>
<td>LST</td>
<td>-0.337</td>
<td>2.583</td>
<td></td>
</tr>
<tr>
<td>TVI</td>
<td>-0.001</td>
<td>3.055</td>
<td></td>
</tr>
<tr>
<td>ALT</td>
<td>-0.013</td>
<td>1.957</td>
<td></td>
</tr>
<tr>
<td>SLOPE</td>
<td>0.006</td>
<td>1.969</td>
<td></td>
</tr>
<tr>
<td>RSR</td>
<td>0.174</td>
<td>4.225</td>
<td></td>
</tr>
<tr>
<td>SMI</td>
<td>-60.451</td>
<td>2.533</td>
<td></td>
</tr>
<tr>
<td>NBR</td>
<td>-0.271</td>
<td>3.660</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Constant</td>
<td>34.880</td>
<td></td>
</tr>
<tr>
<td>ELEVATION</td>
<td>0.001</td>
<td>3.820</td>
<td></td>
</tr>
<tr>
<td>NDVI</td>
<td>-1.008</td>
<td>7.603</td>
<td></td>
</tr>
</tbody>
</table>
### 5.3.1 Model 1

\[ P_1 = \frac{1}{1 + \exp(0.001x_1 - 1.189x_2 - 0.061x_3 - 0.001x_4 - 0.012x_5 + 0.62)} \]  

(5.1)

where, \( P_1 \) is the POP, and \( x_1, x_2, x_3, x_4, \) and \( x_5 \), are respectively elevation, NDVI, LST, TVI, and ALT. In the study of Wang (2017), Elevation, SMI, NDVI and Albedo were involved in the model of POP. In this model 5.1, SMI and albedo were excluded and LST, TVI and ALT were included because these three variables had higher correlations with \( \ln((1-POP)/POP) \) than SMI and also albedo was highly correlated with elevation and TVI. The POP in DTA ranges from 12.89% to 55.99%. The high probabilities exist in the east and northeast parts of the study areas, while the low probabilities are noticed in the south and west parts (Figure 5.21).
5.3.2 Model 2

\[
P_2 = \frac{1}{1 + \exp \left( 0.001x_1 - 1.197x_2 - 0.065x_3 - 0.001x_4 - 0.013x_5 \ldots + 0.005x_6 + 0.087 \right)} (5.2)
\]

where, \( P_2 \) is the POP, and \( x_1, x_2, x_3, x_4, x_5, \) and \( x_6 \) are respectively elevation, NDVI, LST, TVI, ALT, and slope. In addition to the driving factors in model 5.1, slope was added into the model 5.2. The POP in the DTA ranges from 12.82% to 58.74%. The high POP values were found in the east and northeast parts and northwest corner of the study area and the low POP values existed in the south (Figure 5.22).
5.3.3 Model 3

\[ P_3 = \frac{1}{(1 + \exp(0.001x_1 - 1.544x_2 - 0.061x_3 - 0.001x_4 - 0.014x_5 \ldots + 0.007x_6 + 0.107x_7 + 0.098 \ldots))} \]  \hspace{1cm} (5.3)

where, \( P_3 \) is the POP and \( x_1, x_2, x_3, x_4, x_5, x_6, \) and \( x_7 \) are respectively elevation, NDVI, LST, TVI, ALT, slope, and RSR. Compared with the driving factors in the model 5.2, RSR was added into the model 5.3. The POP in DTA areas varied from 20.12% to 58.96%. The high POP values were also found in the east and northeast parts and the northwest corner of the study areas, while the low values existed in the south (Figure 5.23).
Figure 5.23: The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 3.

5.3.4 Model 4

\[ P_4 = \frac{1}{(1 + \exp(0.001x_1 - 1.068x_2 - 0.354x_3 - 0.001x_4 - 0.014x_5 \ldots + 0.005x_6 + 0.139x_7 - 65.378x_8 + 51.012))} \]  

(5.4)

where, \( P_4 \) is the POP and \( x_1, x_2, x_3, x_4, x_5, x_6, x_7, \) and \( x_8 \) are respectively elevation, NDVI, LST, TVI, ALT, slope, RSR, and SMI. In addition to the driving factors in the model 5.3, SMI was introduced into the model 5.4. The POP values were found ranging from 7.49\% to 63.02\%. The POP values had similar spatial distribution with those from the models 1, 2, and 3, but the overall probabilities decreased (Figure 5.24).
Figure 5.24: The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 4.

5.3.5 Model 5

\[
P5 = \frac{1}{(1 + \exp(-0.001x_1 - 1.175x_2 - 0.337x_3 - 0.001x_4 - 0.013x_5 \ldots))}
\]  

(5.5)

where, P5 is the POP and x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, and x_9 are respectively elevation, NDVI, LST, TVI, ALT, slope, RSR, SMI, and NBR. Compared with the driving factors in the model 5.4, NBR was added into the model 5.5. The POP values had a range of from 9.94% to 83.46%. The POP values and their distribution were similar to those from the model 4 (Figure 5.25).
Figure 5.25: The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 5.

5.3.6 Model 6

\[
P_6 = \frac{1}{1 + \text{Exp} \left( 0.001x_1 - 1.008x_2 - 0.255x_3 - 0.001x_4 - 0.015x_5 + 0.007x_6 \ldots + 0.155x_7 - 44.821x_8 - 0.438x_9 + 0.644x_{10} + 47.103 \right)}
\] (5.6)

where, \( P_6 \) is the POP and \( x_1, x_2, \ldots, x_{10} \) are respectively elevation, NDVI, LST, TVI, ALT, slope, RSR, SMI, NBR, and NDSI. In addition to those in the model 5.5, NDSI was added in the model 5.6 due to its relatively high correlation with \( \ln((1-\text{POP})/\text{POP}) \). All the driving factors were statistically significantly correlated with \( \ln((1-\text{POP})/\text{POP}) \). Out of ten involved driving factors, in addition to elevation, SMI and NDVI used in the model presented
by Wang (2017), seven new predictors were selected, which indicated that the new predictors that also significantly contributed the improvement of predicting POP. The albedo utilized in the model of Wang (2017) was not included in the model 5.6 because albedo is highly correlated with TVI. The POP values varied from 9.91% to 83.48%. The overall high POP values occurred in the east, northeast and north parts of the study areas and the low values took place in the south, southwest and west parts. The spatial patterns were consistent with those from all other models (Figure 5.26).

![Permafrost Presence in Donnelly Training Area, Alaska](image)

**Figure 5.26:** The spatial distribution of permafrost probabilities in Donnelly Training Area, Alaska based on the model 6.
5.4 Mapping of Active Layer Thickness

Pearson’s correlation coefficients between ALT and most of the factors, including elevation, SMI, albedo, LST, NDBI, slope, SR, SAVI, EVI, NBR, TVI, and VARI, were statistically significantly different from zero at the significant level of 0.05, indicating potentially significant contributions to prediction of ALT. The predictors that are not significantly correlated with ALT consist of NDVI, NDWI, NDSI, aspect and RSR, implying that these five variables have no statistically significant effects to improving the prediction of ALT. Also, Spearman's Rank correlation coefficients were calculated, there are not much differences from Pearson correlations. The albedo has the highest positive correlation (i.e. 0.36) with ALT (Table 5.2), then TVI, elevation, LST, SMI, NBR, EVI, NDBI, slope, VARI, SAVI, and SR.

Because Ln((1-POP)/POP) is highly correlated with and affected by ALT, mapping ALT was also conducted in this study by a linear model with stepwise regression method and VIF was selected equal to 10. The results of the models were summarized in Table 5.6. A total of eight models were obtained and the p-values of F statistics for all the models were smaller than the significant level of 0.05. The regression coefficients of the models with VIF values of the selected predictors were listed in Table 5.7.
Table 5.6: Summary of models obtained for prediction of active layer thickness using stepwise regression with VIF equal to 10.

<table>
<thead>
<tr>
<th>Model</th>
<th>$R^2$</th>
<th>Adjusted $R^2$</th>
<th>Sig. F</th>
<th>Independent variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.457</td>
<td>0.449</td>
<td>0.000</td>
<td>ALBEDO, SMI, NBR</td>
</tr>
<tr>
<td>2</td>
<td>0.480</td>
<td>0.479</td>
<td>0.000</td>
<td>ALBEDO, SMI, NBR, ELEVATION</td>
</tr>
<tr>
<td>3</td>
<td>0.497</td>
<td>0.481</td>
<td>0.002</td>
<td>ALBEDO, SMI, NBR, ELEVATION, SAVI</td>
</tr>
<tr>
<td>4</td>
<td>0.504</td>
<td>0.503</td>
<td>0.001</td>
<td>ALBEDO, SMI, NBR, ELEVATION, SAVI, NDBI</td>
</tr>
<tr>
<td>5</td>
<td>0.501</td>
<td>0.500</td>
<td>0.023</td>
<td>ALBEDO, SMI, NBR, ELEVATION, SAVI, NDBI, SLOPE</td>
</tr>
<tr>
<td>6</td>
<td>0.507</td>
<td>0.506</td>
<td>0.041</td>
<td>SMI, ELEVATION, SAVI, NDBI, SLOPE, LST, SR, EVI</td>
</tr>
<tr>
<td>7</td>
<td>0.503</td>
<td>0.502</td>
<td>0.023</td>
<td>SMI, ELEVATION, SAVI, NDBI, SLOPE, LST, SR, EVI, VARI</td>
</tr>
<tr>
<td>8</td>
<td>0.519</td>
<td>0.519</td>
<td>0.036</td>
<td>SMI, ELEVATION, SAVI, NDBI, SLOPE, LST, SR, EVI, VARI, TVI</td>
</tr>
</tbody>
</table>


Table 5.7: The regression coefficients and VIF values of the models for prediction of active layer thickness using stepwise regression.

<table>
<thead>
<tr>
<th>Model</th>
<th>Variables</th>
<th>Coefficients</th>
<th>Variation Inflation Factor (VIF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Constant</td>
<td>-293.730</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ALBEDO</td>
<td>31.376</td>
<td>1.392</td>
</tr>
<tr>
<td></td>
<td>SMI</td>
<td>393.992</td>
<td>1.392</td>
</tr>
<tr>
<td></td>
<td>NBR</td>
<td>-11.499</td>
<td>1.456</td>
</tr>
<tr>
<td>2</td>
<td>Constant</td>
<td>-314.848</td>
<td>1.396</td>
</tr>
<tr>
<td></td>
<td>ALBEDO</td>
<td>17.635</td>
<td>1.422</td>
</tr>
<tr>
<td></td>
<td>SMI</td>
<td>418.276</td>
<td>1.022</td>
</tr>
<tr>
<td></td>
<td>NBR</td>
<td>-13.226</td>
<td>1.33</td>
</tr>
<tr>
<td></td>
<td>ELEVATION</td>
<td>0.007</td>
<td>1.434</td>
</tr>
<tr>
<td>3</td>
<td>Constant</td>
<td>-332.711</td>
<td>1.424</td>
</tr>
<tr>
<td></td>
<td>ALBEDO</td>
<td>14.245</td>
<td>1.475</td>
</tr>
<tr>
<td></td>
<td>SMI</td>
<td>437.824</td>
<td>1.518</td>
</tr>
<tr>
<td></td>
<td>NBR</td>
<td>-15.338</td>
<td>1.332</td>
</tr>
<tr>
<td></td>
<td>ELEVATION</td>
<td>0.007</td>
<td>1.482</td>
</tr>
<tr>
<td></td>
<td>SAVI</td>
<td>7.771</td>
<td>1.522</td>
</tr>
<tr>
<td>4</td>
<td>Constant</td>
<td>-339.582</td>
<td>1.665</td>
</tr>
<tr>
<td></td>
<td>ALBEDO</td>
<td>18.621</td>
<td>2.282</td>
</tr>
<tr>
<td></td>
<td>SMI</td>
<td>441.934</td>
<td>1.605</td>
</tr>
<tr>
<td></td>
<td>NBR</td>
<td>9.445</td>
<td>1.657</td>
</tr>
<tr>
<td></td>
<td>ELEVATION</td>
<td>0.006</td>
<td>1.505</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>SAVI</td>
<td>17.299</td>
<td>1.616</td>
<td></td>
</tr>
<tr>
<td>NDBI</td>
<td>97.482</td>
<td>1.741</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-326.776</td>
<td>4.442</td>
<td></td>
</tr>
<tr>
<td>ALBEDO</td>
<td>16.414</td>
<td>3.461</td>
<td></td>
</tr>
<tr>
<td>SMI</td>
<td>425.700</td>
<td>2.800</td>
<td></td>
</tr>
<tr>
<td>NBR</td>
<td>11.104</td>
<td>2.302</td>
<td></td>
</tr>
<tr>
<td>ELEVATION</td>
<td>0.008</td>
<td>2.303</td>
<td></td>
</tr>
<tr>
<td>SAVI</td>
<td>17.245</td>
<td>1.825</td>
<td></td>
</tr>
<tr>
<td>NDBI</td>
<td>101.924</td>
<td>1.796</td>
<td></td>
</tr>
<tr>
<td>SLOPE</td>
<td>-0.108</td>
<td>4.492</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-335.829</td>
<td>3.505</td>
<td></td>
</tr>
<tr>
<td>SMI</td>
<td>13.442</td>
<td>2.822</td>
<td></td>
</tr>
<tr>
<td>ELEVATION</td>
<td>438.239</td>
<td>1.671</td>
<td></td>
</tr>
<tr>
<td>SAVI</td>
<td>16.518</td>
<td>3.287</td>
<td></td>
</tr>
<tr>
<td>NDBI</td>
<td>0.006</td>
<td>2.362</td>
<td></td>
</tr>
<tr>
<td>SLOPE</td>
<td>23.927</td>
<td>3.059</td>
<td></td>
</tr>
<tr>
<td>LST</td>
<td>143.672</td>
<td>1.892</td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>4.156</td>
<td>5.213</td>
<td></td>
</tr>
<tr>
<td>EVI</td>
<td>-18.449</td>
<td>3.607</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-355.217</td>
<td>3.137</td>
<td></td>
</tr>
<tr>
<td>SMI</td>
<td>7.526</td>
<td>1.688</td>
<td></td>
</tr>
<tr>
<td>ELEVATION</td>
<td>460.495</td>
<td>3.830</td>
<td></td>
</tr>
<tr>
<td>Feature</td>
<td>Coefficient 1</td>
<td>Coefficient 2</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>SAVI</td>
<td>14.942</td>
<td>3.598</td>
<td></td>
</tr>
<tr>
<td>NDBI</td>
<td>0.006</td>
<td>2.196</td>
<td></td>
</tr>
<tr>
<td>SLOPE</td>
<td>38.682</td>
<td>1.870</td>
<td></td>
</tr>
<tr>
<td>LST</td>
<td>147.321</td>
<td>4.675</td>
<td></td>
</tr>
<tr>
<td>SR</td>
<td>1.898</td>
<td>3.534</td>
<td></td>
</tr>
<tr>
<td>EVI</td>
<td>-20.584</td>
<td>3.136</td>
<td></td>
</tr>
<tr>
<td>VARI</td>
<td>5.494</td>
<td>1.623</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Feature</th>
<th>Coefficient 1</th>
<th>Coefficient 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>-369.520</td>
<td>2.285</td>
</tr>
<tr>
<td>SMI</td>
<td>5.833</td>
<td>2.303</td>
</tr>
<tr>
<td>ELEVATION</td>
<td>479.614</td>
<td>3.445</td>
</tr>
<tr>
<td>SAVI</td>
<td>16.037</td>
<td>5.678</td>
</tr>
<tr>
<td>NDBI</td>
<td>0.007</td>
<td>2.398</td>
</tr>
<tr>
<td>SLOPE</td>
<td>34.733</td>
<td>4.678</td>
</tr>
<tr>
<td>LST</td>
<td>139.789</td>
<td>1.990</td>
</tr>
<tr>
<td>SR</td>
<td>1.926</td>
<td>3.568</td>
</tr>
<tr>
<td>EVI</td>
<td>-22.981</td>
<td>9.887</td>
</tr>
<tr>
<td>VSRI</td>
<td>3.364</td>
<td>3.220</td>
</tr>
<tr>
<td>TVI</td>
<td>-7.216</td>
<td>0.000</td>
</tr>
</tbody>
</table>

A total of eight models were obtained and shown as follows:

ALT = (31.376x_1 + 393.992x_2 − 11.499x_3 − 293.730)          \tag{5.7}

ALT = (17.635x_1 + 418.276x_2 − 13.226x_3 + 0.007x_4 − 314.848) \tag{5.8}
ALT = (14.245x_1 + 437.824x_2 - 15.338x_3 + 0.007x_4 + 7.771x_5 - 332.711) \quad (5.9) \\
ALT = \left( \frac{18.621x_1 + 441.934x_2 + 9.445x_3 + 0.006x_4 + 17.299x_5 + \ldots}{97.482x_6 - 339.582} \right) \quad (5.10) \\
ALT = \left( 19.424x_1 + 441.057x_2 + 8.825x_3 + 0.006x_4 + 12.642x_5 \ldots \right) \quad (5.11) \\
ALT = \left( 13.442x_1 + 438.239x_2 + 16.518x_3 + 0.006x_4 + 23.927x_5 \ldots \right) \quad (5.12) \\
ALT = \left( 7.526x_1 + 460.495x_2 + 14.942x_3 + 0.006x_4 + 38.682x_5 \ldots \right) \quad (5.13) \\
ALT = \left( 5.833x_1 + 479.614x_2 + 16.037x_3 + 0.007x_4 + 34.733x_5 + 139.789x_6 \ldots \right) \quad (5.14)

where, ALT is the active layer thickness and \(x_1, x_2, x_3, x_4, x_5, x_6, \) and \(x_7\) respectively represent albedo, SMI, NBR, elevation, SAVI, NDBI, and slope for equations 5.7 to 5.11, and \(x_1, x_2, x_3, x_4, x_5, x_7, x_8, x_9, and x_{10}\) respectively represent SMI, elevation, SAVI, NDBI, slope, LST, SR, EVI, VARI, and TVI for equations 5.12 to 5.14. From model 5.7 to model 5.14, all the independent variables that had significant contributions to the dynamics of ALT were selected. In model 5.7, albedo, SMI and NBR were first selected. The TVI is more correlated with ALT than SMI and NBR but TVI was not chosen in model 5.7 because TVI is highly correlated with albedo and NBR.

The range of ALT values varied by model (Figure 5.27), from 4.48 cm to 32.26 cm for model 5.7, from 5.24 cm to 38.75 cm for model 5.8, from 6.23 cm to 41.37 cm for model 5.9, from 2.30 cm to 52.64 cm for model 5.10, from 4.31 cm to 55.37 cm for model 5.11, from 3.89 cm to 59.89 cm for model 5.12, 10.58 cm to 64.98 cm for model 5.13 and from 7.98 cm to 69.56 cm for model 5.14. The spatial distributions of the ALT values obtained from the models were similar to each other (Figure 5.27). The maximum ALT increases from model 5.7 (i.e. max 32.26 cm) to model 5.11 (max 69.56 cm).
cm) to model 5.14 (i.e. max 69.56 cm). The minimum ALT varies from model to model. The values of ALT were greater in the southern, southwestern and western in DTA, indicating greater thickness of active layer mainly due to mountainous, steep and sparsely vegetated areas where seasonal thawing of ice underlain by permafrost took place and thus the areas were more vulnerable for degradation of permafrost. On the other hand, the values of ALT were comparatively smaller in the eastern, northeastern, northern, and northwestern in DTA where the very flat lands are present. That is, the probabilities of permafrost presence were smaller in the southern, southwestern, and western DTA, and greater in in the eastern, northeastern, northern, and northwestern DTA.
Figure 5.27: The spatial distributions of Active Layer Thickness estimates obtained by model 5.7 to model 5.14 in the Donnelly Training Area, Alaska.
Due to less availability of field data it is hard to come up with solid results of POP and ALT estimation validation. For the validation of POP estimates, a total of 414 field observations were used. But, most of the field data were located in the South and Southwest parts of DTA. Even though, the estimates of POP from the models were compared with the field observations and relative root mean square error (RMSE) was calculated.

First of all, based on the available field observation data of POP in DTA, the POP map produced using equation 4.24 from the study of Wang (2017), the obtained relative RMSE is 4.19%. Then, in this study the obtained six models were used to generate the estimates of 414 sample locations using each of the models and the estimates were compared with the field observations. All the models led to similar relative RMSE (Table 5.8). Model 5.6 that had the smaller relative RMSE (i.e. 4.59%) than other models. So, model 5.6 is considered a best model.

For the validation of ALT estimates, a total of 500 points were randomly sampled and their reference values were extracted from the 2010-2019 ALT map developed by Geophysical Institute Permafrost Lab (GIPL, 2017). The ALT estimates were generated using the models developed in this study. Although the estimates were the average values of ALT from Jan 2014 to December 2018, the time period fell in the time interval of the reference values. Comparing the estimates with reference values resulted in the relative
RMSE ranging from 2.39% to 2.54% (Table 5.8). The relative RMSE values were not significantly different from each other among the models. Relatively, the model 5.14 had the best performance.

*Table 5.8: Relative RMSE of POP and ALT estimates from the developed models.*

<table>
<thead>
<tr>
<th>Probability of permafrost presence</th>
<th>Active layer thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Relative RMSE (%)</td>
</tr>
<tr>
<td>MODEL 5.1</td>
<td>5.93</td>
</tr>
<tr>
<td>MODEL 5.2</td>
<td>5.92</td>
</tr>
<tr>
<td>MODEL 5.3</td>
<td>4.88</td>
</tr>
<tr>
<td>MODEL 5.4</td>
<td>4.79</td>
</tr>
<tr>
<td>MODEL 5.5</td>
<td>4.62</td>
</tr>
<tr>
<td>MODEL 5.6</td>
<td>4.59</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Akaike information criterion is another way to validate the produced (Table 5.9). The best model for presence of permafrost in DTA is model 5.6 with AIC value -884.26. This model 5.6 has the predictors including elevation, NDVI, LST, TVI, ALT, SLOPE, RSR, SMI, NBR, and NDSI. Relatively, the model performance becomes poorer from model 5.6 to model 5.1. This means that addition of more predictors improve the performance of the model. The standard model developed by Wang (2017) using elevation, SMI, NDVI, and Albedo leads to the poorest performance in DTA with AIC value of -819.18 (Table 5.9).
The performance of ALT models was also assessed using AIC (Table 5.10). The model 5.14 gives best performance with the smallest AIC value of 655.23 and uses SMI, elevation, SAVI, NDBI, slope, LST, SR, EVI, VARI, and TVI as the model predictors. The model performance becomes better with the addition of the predictors. Model 5.7 has the poorest performance with the greatest AIC values of 917.37, in which albedo, SMI, and NBR were used.
Table 5.10: Akaike Information Criterion (AIC) for Active Layer Thickness (ALT) in DTA

<table>
<thead>
<tr>
<th>Model</th>
<th>N</th>
<th>K</th>
<th>ESS</th>
<th>AIC Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL 5.7</td>
<td>500</td>
<td>3</td>
<td>3192.32</td>
<td>917.37</td>
</tr>
<tr>
<td>MODEL 5.8</td>
<td>500</td>
<td>4</td>
<td>3161.40</td>
<td>913.42</td>
</tr>
<tr>
<td>MODEL 5.9</td>
<td>500</td>
<td>5</td>
<td>3090.77</td>
<td>908.12</td>
</tr>
<tr>
<td>MODEL 5.10</td>
<td>500</td>
<td>6</td>
<td>2994.84</td>
<td>899.70</td>
</tr>
<tr>
<td>MODEL 5.11</td>
<td>500</td>
<td>7</td>
<td>2943.72</td>
<td>860.96</td>
</tr>
<tr>
<td>MODEL 5.12</td>
<td>500</td>
<td>8</td>
<td>2906.45</td>
<td>851.11</td>
</tr>
<tr>
<td>MODEL 5.13</td>
<td>500</td>
<td>9</td>
<td>2881.83</td>
<td>789.15</td>
</tr>
<tr>
<td>MODEL 5.14</td>
<td>500</td>
<td>10</td>
<td>2874.28</td>
<td>655.23</td>
</tr>
</tbody>
</table>
CHAPTER 6
DISCUSSION

This study aims to develop a method of mapping permafrost using optical images and identify the major factors that affect the mapping accuracy. There are two research questions to be answered.

Research Question 1: What are the important predictors that significantly improve mapping the presence of permafrost in DTA?

Research Question 2: How does the permafrost vary spatially in the study area?

6.1 Performance of logistic model with original predictors

The logistic model proposed by Wang (2017) and Yaya et al. (2018) for Qinghai-Tibet Plateau was selected as a basic model in this study. According to Yaya and his colleagues, the model includes the inputs or primary driving factors that significantly influence \( \ln((1-\text{POP})/\text{POP}) \), including elevation, SMI, NDVI, and albedo. The overall accuracy of this model by using the predictors was 85% in Qinghai-Tibet Plateau. Applying the model and the same predictors to predict the POP in DTA led to more accurate estimates with relative RMSE of 4.19%. These factors elevation, SMI, NDVI, and albedo were significantly correlated with \( \ln((1-\text{POP})/\text{POP}) \) in DTA. The elevation had a highest and negative correlation with \( \ln((1-\text{POP})/\text{POP}) \), that is, positively correlated with POP, indicating that the areas with higher elevation had higher probability of permafrost presence. The NDVI had a positive correlation with \( \ln((1-\text{POP})/\text{POP}) \), implying a negative correlation with POP and that the dense the vegetation canopy, the less the likelihood of permafrost existing. The SMI was a combination of NDVI and LST, which might have led to the duplication of information that was ignored in the model of Wang (2017). In fact, in this study, the first five predictors including elevation, NDVI, LST, TVI and ALT were
selected in the model (Table 5.4) of POP. This implied that SMI was not important to improve the prediction of POP due to its information duplication with NDVI and LST. At the same time, albedo was also not selected mainly because both NDVI and TVI were derived by the bands that were involved in albedo (band 1, band 3, band 4, band 5 and band 7). This explained that LST, TVI and ALT were more important than SMI and albedo.

6.2 Contribution of other predictors for predicting probability of permafrost

In this study, in addition to the original factors (elevation, SMI, NDVI, and albedo), other fourteen predictors including LST, NDWI, NDSI, NDBI, Slope, Aspect, SR, SAVI, EVI, NDR, TVI, RSR, VARI and ALT were added to explore improving the prediction of POP. The reasons to add the predictors included 1) Permafrost presence and its dynamics were determined by soil properties, land surface temperature, vegetation, water, soil moisture, burning and snow; and 2) whether or not there were other factors that contributed more to the prediction of POP than the original factors (elevation, SMI, NDVI, and albedo). It was found that all the added predictors except EVI and aspect were statistically significantly correlated with \(\ln \left(1 - \text{POP}/\text{POP}\right)\).

Compared with SMI and albedo, the LST, TVI and ALT contributed more to improving the prediction of POP. The logistic stepwise regression led to more predictors involved in the POP models. As more predictors were added, the adjusted R square value increased. Based on the validation of using the field observations, the model 5.6 led to the smallest relative RMSE. In this model, the selected predictors consisted of elevation, NDVI, LST, TVI, ALT, slope, RSR, SMI, NBR, and NDSI. The contributions of all the selected independent variables to the prediction of POP were statistically significant with VIF values smaller than 10. The predictors especially the image derived indices characterized the topographic features, soil properties, land surface temperature, vegetation, water, burning and snow, and thus made contributions to the
prediction of POP. The performance of the modified models was tested with AIC and the model with the small AIC value gives the best performance. Among the modified models, the model 5.6 is the best model with the smallest AIC values.

6.3 Mapping active layer thickness

The ALT has a correlation coefficient of 0.58 with ln((1-POP)/POP), implying a negative correlation with POP. The relatively high correlation implied that the thicker the ALT, the lower the POP, which thus reasonably explained the relationship of ALT with POP. Due to the importance of ALT, in this study ALT was also mapped using a linear model with eight models obtained. Except for NDVI, NDWI, NDSI, aspect and RSR, all other factors are significantly correlated with ALT. The albedo, SMI and NBR were first selected in model 5.7, and SMI, elevation, SAVI, NDBI, slope, LST, SR, EVI, VARI, and TVI were finally added into the model 5.14 that has the smallest relative RMSE and AIC values. The LST that greatly affected ALT was selected. The independent variables accounted for the characteristics of topography, soil, land surface temperature, vegetation, and water, and thus contributed to the prediction of ALT.

The ALT is a main indicator for permafrost presence because it is the layer that undergoes seasonal thawing and thus ALT varies seasonally (Harris et al., 1988). The ALT is affected by LST and is a critical component in prediction of climate change which is associated with permafrost (Nelson et al., 1998; Zhang et al., 2012). The increase of LST over the last 30 years indicates an increase in permafrost temperatures of 0.5 to 3 °C (Osterkamp, 2005). The increased temperature might have increased the ALT and degraded permafrost (Zhang et al., 2005; Jorgenson et al., 2010). Previous field observation-based studies reported the ALT in several sites in Alaska ranged from 30 to 90 cm (Bockheim, 2007). But, the field observation-based spatial modeling study by Nelson et al. (1997) showed that the ALT ranged to 0 to more
than 70 cm in the heterogeneous types of landscape in the north central Alaska. In this study, the ALT was up to about 66 cm. The study of Hinkel and Nelson (2003) showed that the ALT varied from 20 to 120 cm from 1995 to 2000 at seven sites in northern Alaska. Their study also reported the environmental controls of ALT were dependent on the topographic features, vegetation types, and soil properties. In this study, the maximum depth of ALT varied from about 32 cm to 66 cm depending on the models and the best performance model 5.14 led to the ALT maximum depth of about 70 cm. The maximum ALT value was slightly smaller than those in the previous studies. There are 8 models of ALT with different model factors. Addition of model factors the model performance proved by the relative RMSE and values. Model 5.14 had the maximum number of the factors with the smallest AIC values.

6.4 Spatial distribution of Permafrost and active layer thickness

According to Jorgenson et al. (2008), discontinuous type of permafrost exists in DTA. In this study, it was found that the probability for this discontinuous type of permafrost presence varied over the study area. The east, northeast, north and northwest parts were found to have higher POP values, and the central and south to southwest parts had smaller probabilities of permafrost presence. The spatial distribution of ALT was inversely consistent with that of POP. That is, the ALT was higher in the central, south, and southwest parts due to complex topographic, higher LST and sparse vegetation canopy, and lower in the east, northeast, north and northwest parts due to the lower LST, forests and wetlands.

6.5 Limitation

The major limitation of this study was the availability of field data in terms of the time for the field observations collected and their spatial distribution. This made the validation of the model results less possible. Although some of the field data were used, the validation of the
obtained results was limited due to different time at which the used field observations were collected. Another limitation came from the medium spatial resolution (30 m by 30 m) of the used Landsat images. In addition, a limited number of remote sensing spectral variables were explored for their contributions to the prediction of POP and ALT. It is possible to leave some factors that significantly affect POP and ALT out of the examination in DTA.
CHAPTER 7

CONCLUSIONS

Permafrost occupies about a quarter of the northern hemisphere land, representing a massive amount of 25.5 million ha. Due to the cost and complex technologies required by processing InSAR images, this paper explored the use of freely downloadable Landsat satellite images for mapping POP and ALT in DTA, Alaska. Landsat 8 OLI/TIRS images from January 2014 to December 2018 were utilized and the factors that affect POP and ALT and their dynamics were analyzed based on a logistic model and a linear regression model, respectively. A total of seventeen factors were utilized and selected using stepwise regression with VIF of 10.

The results showed that the POP in DTA was significantly affected by all the factors except aspect and EVI. The factor that was most correlated with ln((1-POP)/POP) was elevation, then NDVI, albedo, ALT, LST, NDWI, NDSI, slope, TVI, RSR, SMI, NDBI, SR, SAVI, NBR and VARI. The elevation, NDVI, LST, TVI, ALT, SLOPE, RSR, SMI, NBR, and NDSI were finally chosen in the best model 5.6. The previously used albedo was excluded in the final model. This implied that the albedo was not critical to the prediction of POP. In addition to elevation, NDVI and SMI, other predictors including LST, TVI, ALT, SLOPE, RSR, NBR and NDSI could not be ignored in the prediction of POP because they characterized the properties of topography, soil, vegetation, land surface temperature, water, building and burning. The model generated reasonable spatial distribution of POP in which POP had greater values in the east, northeast, north and northwest parts and smaller in the south and southwest parts.

Except for NDVI, NDWI, NDSI, aspect and RSR, moreover, all other predictors are significantly correlated with ALT. SMI, ELEVATION, SAVI, NDBI, SLOPE, LST, SR, EVI,
VARI and TVI were finally selected in the best model 5.14. The ALT highly varied over the study area and its spatial patterns were inversely consistent with those of POP.

The results are essential for the governments, policymakers, and other concerned stakeholders to estimate the degradation of permafrost in DTA and minimize the risk of policy decision-making for land use management and planning. This study will help to understand the global climate change, changing ecosystem, increasing concentration in the atmosphere, and human activity induced disturbance.
REFERENCES


Alaska Climate Research Center (ACRC). 1999. Fairbanks Alaska Climatology, Geophysical Institute, Univ. of Alaska Fairbanks, AK (http://climate.gi.alaska.edu).


high-resolution satellite imagery and ground measurements. *Remote sensing of environment, 80*(1), 165-184.


John Wiley & Sons.
degradation and ecological changes associated with a warming climate in central

Marchenko, S. 2010. Resilience and vulnerability of permafrost to climate
change. *Canadian Journal of Forest Research, 40*(7), 1219-1236.

Jorgenson, M. T., Shur, Y. L., & Pullman, E. R. 2006. Abrupt increase in permafrost degradation

Jorgenson, M. T., Yoshikawa, K., Kanveskiy, M., Shur, Y. L., Romanovsky, V., Marchenko, S.,
Engineering. *University of Alaska, Fairbanks*.

sensing: A study of Vadodara urban area, Gujarat. *International Journal of Geology,
Earth and Environmental Sciences, 2*(1), 123-130.

Inst. d. Univ.

periglacial processes, 19*(2), 107-136.


Mulla, D. J. 2013. Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems engineering*, 114(4), 358-371.


Smith, R. B. 2010. The heat budget of the earth’s surface deduced from space. *Yale University Center for Earth Observation: New Haven, CT, USA*.


Appendix 1: Circum-Arctic Map of Permafrost and Ground-Ice Conditions, Version 2. [Indicate subset used]. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center (Brown et al., 2002).
Appendix 2: Brightness temperature for band 10 on left and band 11 on right of 01/30/2014.
VITA

Graduate School
Southern Illinois University

Kiran Thapa
thapakiran668@gmail.com

Tribhuvan University Kathmandu, Nepal
Bachelor of Science, Geology, July 2008

Thesis Paper Title:

Contributions of Optical Remote Sensing to Permafrost Mapping in Donnelly Training Area, Alaska

Major Professor: Dr. Guangxing Wang