

Machine learning as a tool to understand bio-geophysical indicators of landscape sensitivity to permafrost related mass wasting



Kaitlyn Diederichs, Trevor Lantz

School of Environmental Studies – University of Victoria, Victoria, British Columbia, Canada

Peter Morse

Geological Survey of Canada – Natural Resources Canada, Ottawa, Canada

ABSTRACT

In Northwestern Canada the Dempster and Inuvik-to-Tuktoyaktuk (ITH) highways create the only all-season transportation corridor between Dawson City, Yukon and Tuktoyaktuk, Northwest Territories. These roads traverse a variety of landscape types and ecoregions that are underlain by permafrost. As permafrost thaws in response to climate warming, permafrost geohazards such as mass wasting by active layer detachments and retrogressive thaw slumps are impacting road operation and human activity in the region. This paper uses random forest machine learning to investigate the relative importance of topographic, geological, climatological, and ecological variables describing a series of these mapped permafrost thaw features within a 5-kilometer buffer on either side of this highway corridor. This provides a valuable window into the distribution of thaw sensitive sites across the landscape, and the distribution patterns that can be analyzed within their bio-geophysical contexts to identify potentially geohazardous terrains. The design of random forest models to align with desired interpretation goals is also discussed.

RÉSUMÉ

Dans le nord-ouest du Canada, les autoroutes Dempster et Inuvik-Tuktoyaktuk (ITH) créent le seul corridor de transport toutes saisons entre Dawson City, Yukon, et Tuktoyaktuk, NWT. Ces routes traversent une variété de types de paysages et d'écorégions dont le substrat est le pergélisol. À mesure que le pergélisol dégèle en réponse au réchauffement climatique, les géorisques liés au pergélisol, tels que la perte de masse due aux détachements de couches actives et les effondrements rétrogressifs dus au dégel, ont un impact sur l'exploitation routière et l'activité humaine dans la région. Cet article utilise l'apprentissage automatique forestier aléatoire pour étudier l'importance relative des variables topographiques, géologiques, climatologiques et écologiques décrivant une série de ces caractéristiques cartographiées du dégel du pergélisol dans une zone tampon de 5 kilomètres de chaque côté de ce corridor routier. Cela fournit une fenêtre précieuse sur la répartition des sites prédisposés au dégel dans le paysage et sur les modèles de répartition qui peuvent être analysés dans leurs contextes biogéophysiques pour identifier les terrains potentiellement géodangereux. La conception de modèles forestiers aléatoires pour s'aligner sur les objectifs d'interprétation souhaités est discutée aussi.

1 INTRODUCTION

Permafrost – or perennially frozen ground (Harris et al., 1988) – is the foundation of Arctic landscapes, underlying approximately 40% of the Canadian landmass (Derksen et al., 2019), with more extensive distribution in Northern latitudes. Ice-bonded sediments are often ice rich, thus as climate change progresses, trends towards increased permafrost thaw (Åkerman & Johansson, 2008; Biskaborn et al., 2019) will result in landscape changes that include intensification of thaw-driven mass wasting (Kokelj et al., 2021a). This poses a threat to infrastructure (Hjort et al., 2018; Stevens, 2020), the environment (Schuur & Mack, 2018) and the continuation of important cultural activities such as harvesting and travel (Andrews et al., 2016).

Permafrost thaw is a three-dimensional phenomenon, and the connectivity, thickness and surrounding landscape characteristics reflect the rate, directionality (Fisher et al., 2016; McClymont et al., 2013) and impacts of thaw. This paper investigates a machine learning approach for ranking the measurable landscape characteristics most important in assessing the likelihood of mass wasting responses to permafrost thaw in a given area, providing an opportunity for proactive mitigation. Building from the machine learning approach for thaw-susceptibility

modelling advanced by Rudy et al. (2019), we propose that increasing the specificity of machine learning models increases their interpretability and usefulness for understanding landscape scale trends in mass wasting likelihood.

1.1 Study Area

This paper focuses on a 10-kilometer-wide study area centered on the Dempster Highway and Inuvik to Tuktoyaktuk Highway (ITH) corridors in Northern Canada. The Dempster Highway is a 738km long gravel road that stretches from just east of Dawson City, Yukon to Inuvik, NWT. The ITH is a newer 138km extension of the all-season transportation corridor that continues from Inuvik to Tuktoyaktuk, at the coast of the Arctic Ocean. The majority of the Dempster Highway and the entirety of the ITH is within the continuous permafrost zone (Heginbottom et al., 1995). The highway corridor creates a pseudo-transect through a variety of permafrost conditions with variable geologic and climate contexts, landscape types and ecoregions. This diversity in bio-geophysical characteristics allows for a comprehensive glimpse into the factors that control the distribution of mass wasting related responses of the landscape to a changing climate.

1.2 Permafrost Related Geomorphic Features

Within this paper, two types of thaw-driven mass wasting features are investigated, active layer detachments (ALD) and retrogressive thaw slumps (RTS).

ALDs develop most commonly in silty clays with low to medium plasticity in moderately sloped terrain when the base of the active layer detaches from underlying permafrost (Lewkowicz & Harris, 2005b). Detachment produces a headwall, and a scar zone forms as thawed materials slide downslope on a defined shear plane that exposes underlying permafrost (Lewkowicz & Harris, 2005a). Detachment is initiated or propelled by lowered cohesion along the basal shear plane, in turn driven by excessive pore pressures along this plane resulting in a loss of soil strength as shear stresses due to the gravity of the overlying material exceed shear strengths (Harris et al., 1988; Harris & Lewkowicz, 2000). Critical pore-pressures may develop when ice rich active layers or ice lenses thaw over impermeable permafrost either due to high surface temperatures, forest fire activity or meteorological events such as heavy rainfall (Lewkowicz & Harris, 2005a). While ALDs are a category of mass wasting landform in their own right, the scar zone of the ALD can develop into an RTS if degradation continues or deepens (Jorgenson & Osterkamp, 2005).

An RTS forms when ice-rich permafrost is exposed and thaws (Harris et al., 1988). They are initiated by a variety of processes and frequently occur adjacent to lakes and creeks including slumping and subsequent erosion into these adjacent watercourses (Burn & Lewkowicz, 1990), but the scar zone of the ALD can also develop into an RTS if ice-rich permafrost is exposed and thawing continues (Jorgenson & Osterkamp, 2005). Thawing creates a steep headwall as liberated water and sediments run down the thaw face and are transported rapidly downslope as a debris flow (Harris et al., 1988; Ward Jones et al., 2019).

The thawing and wasting processes and retrogressive retreat of the headwall continue until the headwall reaches ice-poor permafrost or there is insufficient transport of thawed materials downslope, and the headwall becomes covered by debris and insulated from further thaw (Burn & Lewkowicz, 1990). RTSs are typically deeper than the surrounding active layer.

1.3 Geomorphic Feature Inventory

Sladen et al. (2022) mapped over 2000 mass wasting landforms throughout the study area (Figure 1) through visual inspection of high-resolution satellite imagery and classified them into types and subtypes of features according to the methodology developed by Sladen et al. (2021). This dataset is comprised of polygon boundaries for each landform, and these provided training data and labels for the models discussed in this paper. Within the context of this paper, we selected the ALD and RTS subtypes of flow features category for comparison due to their similar ranges (as seen in Figure 1) and because of the potential for ALDs to develop into RTSs.

The distribution of these features across the study area is non-uniform, with certain physiographic regions displaying higher concentrations than others, as can be seen in Figure 1. There are 149 ALD features within the inventory. ALDs are present from the Ogilvie Mountains area through to the Anderson Plain, with the highest concentration occurring in the Eagle Lowlands. There are 326 RTS features within the inventory. RTSs are present throughout the study area north of the Klondike Plateau, and generally increase in frequency in the more northern physiographic regions. There is a very low frequency of RTS landforms in the Eagle Lowland region and a very low frequency of ALD landforms in the Tuktoyaktuk coastlands region.

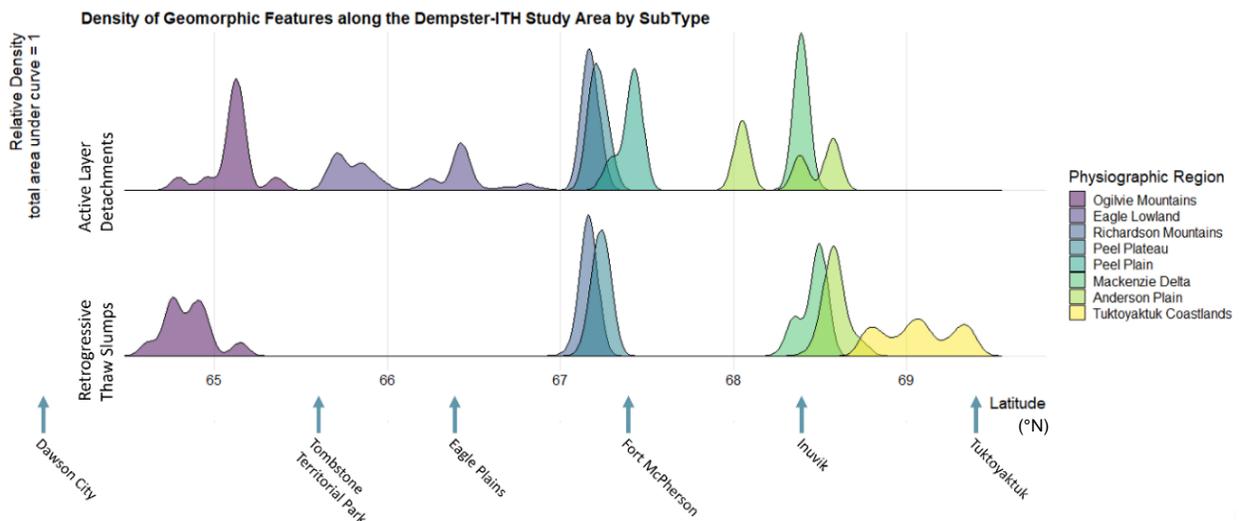


Figure 1: Distribution of observations of each geomorphic feature subtype by physiographic region and latitude. Vertical axis is not to scale between feature types, and instead represents the relative density distributions of a given feature in each physiographic region.

2 DATA ANALYSIS METHODS

2.1 Random Forest

A random forest (RF) model is a type of machine learning classifier built using a series of decision-making classification and regression trees (CART). CARTs describe partitions within the feature space of a data set that allows for the data to be divided into a finite number of classes (Loh, 2011; Prasad et al., 2006). Each CART is grown using supervised learning methods in which data is labelled manually prior to ingestion. Partitions are made on a per-variable basis in a way that maximizes homogeneity within a grouping and heterogeneity between groupings. Diversity of variable precisions and ranges are well accounted for in RF modeling (Strobl et al., 2007), as well as non-linear relationships and highly correlated or intractable variable interactions (Cutler et al., 2007).

A single CART creates an intuitive classifier, meaning it can be easily interpreted by the modeler, however it has the potential to overfit to the original data (known as training data) due to its simplicity (Biau & Scornet, 2016). The RF model solves this issue by creating decision redundancy (Biau & Scornet, 2016; Breiman, 2001) and a diversity of decision pathways. This is done through the use of random subsets of the training data and training variables into unique training sets for each CART in the RF model, which increases the output stability (Cutler et al., 2007; Prasad et al., 2006). Additionally, this “bootstrap sub-setting” allows for unbiased error rate calculations, because the testing data does not need to be separated from the training data when building the model (Prasad et al., 2006). This error rate is called the “out of bag error” and is the overall error when classifying each training point with a model created without that point.

RF models can be interpreted through an analysis of the importance of each variable and through the response of the outcome to each variable. The Variable Importance Plot (VIP) displays the marginal influences of each variable on the model output accuracy, when a given variable is removed from the model. When input variables are uncorrelated, this method gives a good indication of which variables are the most important to the classification of the dataset. When variables are significantly correlated, their individual influence on the model may be artificially diminished relative to the influence of the described feature on the real-world phenomenon being modeled (Wies et al., 2023). It is important to note that the VIP gives an indication of the most important dimensions to the overall model, rather than to each individual decision the model makes.

2.2 Explanatory Variables

We compiled a set of 36 variable inputs, enumerated in Table 1, that describe landscape characteristics related to permafrost, geology, climate, ecology, and geography. These have the potential to influence permafrost thaw sensitivity and were processed and aggregated using the R statistical computing software (R Core Team, 2022).

Slope, elevation and aspect are topographic variables correlated with increased thaw slump activity in permafrost terrains within our study area (Lacelle et al., 2015). We

determined these variables for our study area using the ArcticDEM digital elevation model (10 metre spatial resolution, 4 metre absolute accuracy in any given direction) (Porter & et al., 2022). Elevation was directly extracted, and slope and aspect were calculated using the terrain function from the Terra package in R (Hijmans, 2022). To eliminate a false distance between aspect values on either side of North, we introduced southern (Equation 1) and western (Equation 2) exposure metrics (Brenning & Trombotto, 2006).

$$\begin{aligned} \text{southern exposure} &= -\cos(\text{aspect}) & [1] \\ \text{western exposure} &= -\sin(\text{aspect}) & [2] \end{aligned}$$

Landcover characteristics are correlated with active layer thicknesses (ALT) (Smith et al., 2009) and so were included as a potential mass wasting driver in the RF model, as ALT. They can also impact the fine-scale variability of snow cover due to wind drift and collection (Sturm et al., 2001), further influencing the thermodynamic regime of the soil and the ALT. The dataset used categorizes the likeliest dominant vegetation type for each grid cell. This data is sourced from the Wang et al. (2019) RF model derived classification, which provides a 30-metre resolution gridded dataset of 15 landcover classes for each year from 1984 to 2014. Classification was performed using RF modelling on Landsat surface reflectance data, high resolution imagery and field photography, resulting in predictions for the annual dominant plant functional type in each pixel. The modal landcover for each pixel across the 31-year temporal range was taken for this analysis.

Changes to the ALT are also influenced by occurrences of forest fires (Fisher et al., 2016). Forest fires can also impact the infiltration potential of an area due to vegetation changes, modifying slope responses to rainfall, a potential landslide trigger (Kokelj et al., 2015; Young et al., 2022). The presence of past forest fire activity has been linked to increased rates of mass wasting in areas adjacent to the study area. Fire history data for the study area was sourced from the territorial geomatics databases for both the NWT (NWT Centre for Geomatics, 2019) and the Yukon (Wildland Fire Management - Government of Yukon, 2014) in the form of vector boundaries of fire events. The total area of each fire and the year of occurrence were extracted for each data point.

Surficial geology was sourced from the Quaternary Geology of Canada and Greenland map as prepared by Fulton (1989), available in a vector format. No minimum resolution value is provided, but the dataset includes a smallest polygon element of approximately 20m². While this denotes the smallest bounded area captured by the map, there may be heterogeneity within each zone that is not captured by this data product. The permafrost zones – taken from Heginbottom et al. (1995) – are also provided as vectors and give estimates of areas with high likelihoods of each permafrost density category but may not reflect fine scale conditions.

Ground ice probability data is sourced from the Ground Ice Map of Canada (O’Neill et al., 2022), and provides an estimate of the combined volumetric percentage of excess ice in the top 5 m of permafrost. This data is also available as separated segregated, wedge and relict ice. This ground

ice dataset was built based on surficial materials, permafrost distribution, glacial extents, and modeled paleo vegetation (O'Neill et al., 2019), and the potential prevalence of ice is represented with qualitative descriptors. Absolute accuracy of the model is not available, but it should be noted that validation of the model involved a combination of comparisons to field data on ice presence as well as to the observation of ground ice related geomorphometric features. This indicates that there may be some tangled causality between the presence of the mass wasting sites investigated within this project and the predictions of the ground ice concentrations in the dataset generated with the O'Neill et al. model (2019). That is, the ground ice data being used to predict the presence of mass wasting sites may have been generated using the presence of these same sites. The distance to the Wisconsinan glacial limit of the Laurentide Ice Sheet (Duk-Rodkin, 1999) was also calculated for each location within the glacial extent. A flag value was given to datapoints located beyond (southwest of) the glacial extent, rather than an explicit distance.

RTSs are often instigated by hydraulic erosion at the toe of the slope (Burn & Lewkowicz, 1990), and there are growing concerns about the impact of RTS features on the chemistry and overall ecological health of adjacent streams and lakes (Kokelj et al., 2005, 2013, 2021), indicating that there are frequent slumps in close proximity to water features in the region. The distance to the nearest watercourse or lake can also give an indication of the general soil moisture regime, which affects the pore pressure and resulting shear strength of thawed soils as well as influences ALT and the overall thermodynamic regime of the soil (Clayton et al., 2021).

The final set of variables used in this analysis was a qualitative dataset that divides the study area into a set of three climate zones and a quantitative dataset that includes 19 calculated bioclimatic variables. The climate zones include a tundra climate, a subarctic climate with year-round precipitation and a continental subarctic climate with dry summers (Beck et al., 2018). The bioclimatic dataset consists of a series of calculated metrics that combine precipitation and temperature data from rasterized climate models to create descriptors of local climate behaviour and its temporal variability (Fick & Hijmans, 2017).

Table 1. Explanatory variables used in Random Forest modelling.

Explanatory Variable	count	Data Type (Units)	Source
Elevation	(1)	10 metre raster (m)	PGC Arctic DEM (Porter & et al., 2022)
Slope	(1)	10 metre raster (degrees)	ArcticDEM derivative
Aspect (and south/west exposure metrics)	(2)	10 metre raster (degrees)	ArcticDEM derivative

Landcover	(1)	30 m raster (categorical)	NASA ABoVE (Wang et al., 2019)
Fire History	(2)	vector dataset of event records (fire area, fire)	Territorial Governments (NWT Centre for Geomatics, 2019; Wildland Fire Management - Government of Yukon, 2014)
Surficial Geology	(1)	vector dataset (categorical)	Geological Survey of Canada (Fulton, 1989)
Permafrost Extents	(1)	vector dataset (categorical)	Natural Resources Canada (Heginbottom et al., 1995)
Ground Ice (including bulk, segregated, wedge and relict ice)	(4)	1 kilometre raster dataset (ranked categories of prevalence)	GSC (O'Neill et al., 2022)
Distance to Glacial Limit	(1)	Euclidean distance calculated to vector dataset (m)	GSC (Duk-Rodkin, 1999)
Distance to Watercourses and Lakes	(2)	Euclidean distance calculated to 1:50,000 scale vector dataset (m)	Government of Canada (Natural Resources Canada, 2017)
BioClimatic Variables (full list in Appendix)	(19)	1 kilometre raster of various calculated averages/variances	WorldClim (Fick & Hijmans, 2017)
Climate Zones	(1)	1 kilometre raster (categorical)	Nature: Scientific Data (Beck et al., 2018)

2.3 Training Dataset

Datapoints input into the RF model building algorithm were created as a combination of the polygons outlining identified mass wasting landforms from Sladen et al. (2022) and an equal number of randomly distributed points located within the study area exclusive of these polygons. The randomly distributed points form the centroids of the control set of sampling polygons, each with a circular buffer of an area equal to the median mass wasting polygon from the mass wasting inventory. The 950 total datapoints are labeled with their subtype or as control points for the purposes of classification. The locations of each of these points can be seen in Figure 2.

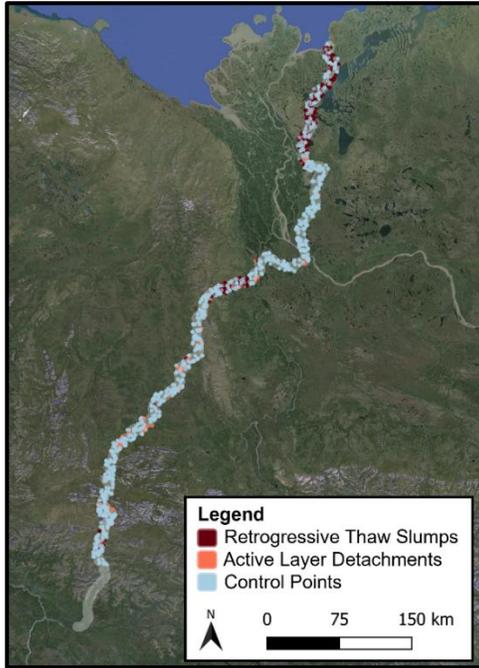


Figure 2: Locations of RTS and ALD features within the study area boundary as well as generated control points for the machine learning training dataset.

For sampling, we used the finest reasonable resolution of each explanatory variable and performed this for each parameter with the *extract* function from the Terra package (Hijmans, 2022). Categorical explanatory variables sourced as vector datasets were sampled based on the proportional inclusion of the categorical polygon within sampling polygon boundaries. Calculations of Euclidean distances between centroids of sampling polygons and vector datasets (glacial limit, watercourses, and lakes) were performed using the *st_distance* function from the Simple Features R package (Pebesma, 2018).

3 RESULTS

Using the training dataset, we constructed a series of five random forest models. Each model contained all the variables listed in Table 1, with different dataset partitions based on the labels of the points. The first two models compared ALD and RTS datapoints to the control set. These models are designed to give an idea of the relative difference between each type of mass wasting affected area and the background landscape. The last three models served to compare these two mass wasting types to each other. Model 3 combined the ALD and RTS occurrence datapoints into one category against the control points. Model 4 directly compared the RTS and ALD datapoints without any control points and Model 5 compared all three classes. The results of these model runs are described in the following sections.

3.1 Active Layer Detachments

There are 149 ALDs within the inventory for which a complete set of variables can be compiled. The random forest model generated from these 149 points and 149 randomly spaced control points had an out of bag error estimation of 8.05%, with 16 false positives and 8 false negatives for the prediction of ALD landforms amongst the 298 training points (Table 3).

Table 3. Confusion matrix for the ALD presence/absence random forest model (Model 1).

		Predicted Class		Class Error
		ALD	Control	
True Class	ALD	141	8	5.37%
	Control	16	133	10.7%

The most important variable in determining the presence of an ALD when compared to the control points was slope, followed by the mean annual temperature (labelled as BIO1 in Figure 3), south exposure, maximum temperature of the warmest month (BIO5), the temperature of the wettest quarter (BIO8), the distance to the nearest watercourse and the mean temperature of the warmest quarter (BIO10), the mean temperature of the driest quarter (BIO9), the mean diurnal temperature range (BIO2) and the annual precipitation (BIO12) of a given datapoint's location. The VIP of this model is shown in Figure 3. There is no prescribed cut-off in mean decrease in accuracy to determine whether a variable is irrelevant to a model, but the top 10 most important variables have been shown in Figure 3 for clarity.

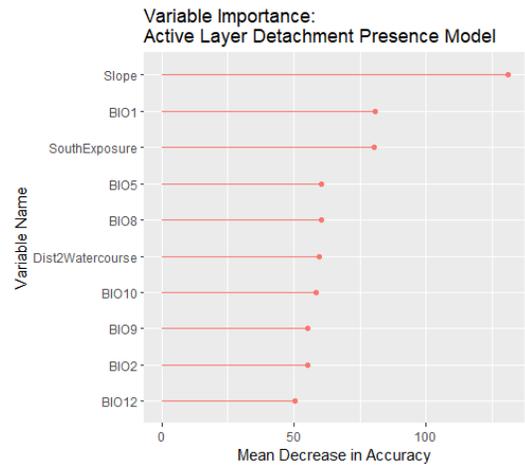


Figure 3: Variable importance plot displaying the 10 most important variables in predicting ALD presence in comparison to control points (Model 1).

3.2 Retrogressive Thaw Slumps

There are 326 RTSs within the inventory for which a complete set of variables can be compiled. The random forest model generated from the 326 slump occurrence

points and the 326 randomly spaced control points had an out of bag error estimation of 6.13%, with 28 false positives and 12 false negatives for the prediction of RTS landforms amongst the 652 training points (Table 4).

Table 4. Confusion matrix for the RTS presence/absence random forest model (Model 2).

		Predicted Class		Class Error
		RTS	Control	
True Class	RTS	314	12	3.68%
	Control	28	298	8.59%

The most important variable in determining the presence of a RTS when compared to the control points was slope, followed by the distance to the nearest lake, the mean diurnal temperature range, south exposure, isothermality (diurnal temperature range/annual temperature range), elevation, the precipitation seasonality, the mean temperature of the warmest quarter, the annual mean temperature, and the annual temperature range.

3.3 Comparison of ALDs and RTSs

There is a total of 475 RTS and ALD sites within the inventory for which a complete set of variables can be compiled. Three models were created with both ALD and RTS points included. Model 3 considers these 475 points to belong to a single category of mass wasting presence in comparison to 475 control points representing an absence of mass wasting sites. Model 4 directly compares the two landform types with no control points. Model 5 compared both types of mass wasting landforms and the control points as individual categories.

Model 3, generated to compare the combined ALD and RTS points to the control points, has an out of bag error estimation of 7.37% with 36 false positive and 34 false negative predictions of mass wasting landform presence across the 950 training points (Table 5). The five most important variables are slope, distance to the nearest lake, south exposure, elevation, and the mean diurnal temperature range respectively.

Table 5. Confusion matrix for the RTS and ALD combined presence/absence random forest model (Model 3).

		Predicted Class		Class Error
		RTS/ALD	Control	
True Class	RTS/ALD	441	34	7.16%
	Control	36	439	7.58%

Model 4 compares only the ALD and RTS points with no control points. The out of bag error estimation is 3.79% with 12 of 149 ALD sites incorrectly classified as RTS sites and 6 of 326 RTS sites incorrectly classified as ALD sites (Table 6). The five most important variables for this model are mean diurnal temperature range, south exposure, slope, annual mean temperature, and the maximum temperature of the warmest month.

Table 6. Confusion matrix for the RTS and ALD comparison random forest model (Model 4).

		Predicted Class		Class Error
		ALD	RTS	
True Class	ALD	137	12	8.05%
	RTS	6	320	1.84%

The final model generated to compare all three categories of points (ALD, RTS and control) has an out of bag error estimate of 8.53%. Table 7 shows the confusion matrix for this model. The random forest model classifying these three categories was most likely to mislabel ALD points. For this model, the five most important variables are slope, distance to the nearest lake, south exposure, elevation, and the mean diurnal temperature range.

Table 7. Confusion matrix for the RTS and ALD comparison random forest model (Model 5).

		Predicted Class			Class Error
		ALD	RTS	Control	
True Class	ALD	122	9	18	18.1%
	RTS	2	304	21	6.75%
	Control	6	26	443	6.74%

4 DISCUSSION

There are three model outputs through which the generated RF models have been interpreted. The first metric for model interpretation is the out of bag error, or prediction accuracy, of each model and the resulting confusion matrix (Tables 4 through 7). Model 1 was built solely to classify ALDs and has a prediction accuracy of 91% while Model 2, built to classify RTSs has a prediction accuracy of 93%. These results are quite similar, particularly given the differences in dataset size, and indicate a very good model fit, with both models displaying higher class accuracy for the landform presence points than the control points. This follows from the assumption that control points taken randomly across the study area will display higher variability in their characteristics than points taken within areas experiencing similar geomorphic processes. The combined presence absence model (Model 3) has an accuracy of 92% and the three-class model (Model 5) has an accuracy of 91%, though the ALD class had a significantly higher misclassification rate than the other classes but is obscured by the lower number of points in this class as compared to the other two. These results overall indicate that the models fit the dataset well, and that the other two metrics of interpretation discussed here can be legitimately used.

The second metric of evaluation that we used in RF interpretation is the variable importance rankings. For models where control points are included (Models 1, 2, 3 and 5), slope is the most important variable in classification. Model 4 directly compares ALD and RTS occurrence without control points, and diurnal temperature range and southern exposure rise above slope in importance, indicating that in that specific decision, the

slope of the ground is important but not as critical as diurnal temperature range and the aspect of the slope. Similarly, in the four models that contain control points in their training datasets (Models 1, 2, 3 and 5), the distance to the nearest waterbody ranked among the most important variables. This was not the case in comparing RTS and ALDs (Model 4), where distance to the nearest watercourse and nearest lake ranked 11th and 12th out of the variable inputs in terms of importance to model accuracy. This difference in variable importance between models allows for an otherwise black box model to be interpreted in a higher granularity due to the specificity of the models. The value added to the interpretability of these models means that landscapes can be parsed for clusters of characteristics that would indicate increased likelihood of one or more geomorphic features developing, independent of their morphological similarities.

The final method of RF interpretation that we used is the model response to individual variables, known as a partial dependence curve. This result shows how a model reacts to variance in one dimension when all others are held steady. For example, in the case of classifying RTSs vs ALDs, Model 4 and 5 are both more likely to choose the RTS class for shallower slopes and the ALD class for steeper slopes. Additionally, across the study area, ALDs are more likely to occur in areas with higher average annual temperatures than RTSs. This method of interpretation can show both real responses to landscape variables or it can indicate the potential for there to be underlying geographic reasons for the link between the occurrence of a mass wasting feature and a variable. It is possible that at a local scale, these features do respond to small scale temperature variation in this way. It is also possible that the apparent presentation of this relationship is an artifact of the geographic distribution of the features within the study area. There is a higher density of ALDs in the warmer southern reaches of the study area, and a higher density of RTSs in the cooler Tuktoyaktuk coastlands near the north. Thus, the individual partial dependence of a variable must be analyzed in concert with other influential variables and explanatory geologic processes in order to truly understand the causality of that specific variable. At the scale of the Dempster-ITH study area, the links between each of the climate variables and mass wasting activity yield classification results of a high accuracy, but it is important to understand that models built only on this training set may not be scalable without an expanded dataset or further analysis of both smaller or larger regions. This is due to a high correlation between latitude and average annual temperature and the nearly linear nature and low longitudinal variation of the study area. This is potentially true for any analysis where variables follow spatial trends and should be kept in mind when generating models over large areas.

Generally, variables with higher spatial resolution proved to be more important in model classification at the landform boundary scale. Due to the fine scale and limited area covered by this dataset, variables that captured fine scale landscape variability ranked quite highly in each model's variable importance results. RF models built for similar purposes using coarser scale feature inventories (such as rasters of frequencies of landform occurrence) or

broader study areas may be better suited to leveraging data of lower resolutions.

The models presented in this paper demonstrate the feasibility of creating a random forest model for the classification of geomorphic feature presence in permafrost zones with great accuracy. They also highlight that a deeper understanding can be gained from interpretation of the model characteristics when generated on very specific subsets of data and classes designed for the investigation of specific relationships between features and landscape descriptors. The multiple configurations of RF models increases the understanding of how each model is making the decisions it outputs, permits a rapid assessment of critical landscape characteristics for different feature occurrences and allows for more confidence in the final classification.

5 FUTURE WORK

This paper presents an analysis of a subset of the available mass wasting feature dataset and outlines three methods for interpreting and fine-tuning RF models based on outputs and behaviour. Further analysis of the other subtypes of geomorphic features present in the Dempster-ITH corridor will complete the set of intra-class variable importance rankings. This creates opportunities for high dimensional analysis of the differences in likelihood of each type of feature in the landscape currently and under future climate scenarios. Additionally, performing this analysis over geographic clusters of features will provide insight into these differences over different landscape presentations and ideally eliminate or quantify the effects of geometric artifacts in the model results.

While not possible for an analysis of the inventory used in this paper, random forest models of the type discussed here could be used with dated mass wasting inventories to create time-series of risk related to vegetation and landcover change as well as climate variations with increasing global temperatures.

6 ACKNOWLEDGMENTS

This work was made possible through the support of the Natural Sciences and Engineering Research Council of Canada (NSERC) (PermafrostNet NETGP 523228 – 18). The Natural Resources Canada contribution number for this paper is 20230412. A huge thank you to Hayley Webster and Jean Hutchinson for their time and insights in editing this paper.

7 REFERENCES

- Åkerman, H. J., & Johansson, M. (2008). Thawing permafrost and thicker active layers in sub-arctic Sweden. *Permafrost and Periglacial Processes*, 19(3), 279–292. <https://doi.org/10.1002/ppp.626>
- Andrews, T. D., Kokelj, S. V., MacKay, G., Buysse, J., Kritsch, I., Andre, A., & Lantz, T. (2016). Permafrost Thaw and Aboriginal Cultural Landscapes in the Gwich'in Region, Canada. *The Journal of Preservation Technology*, 47(1), 15–22.

- Beck, H. E., Zimmermann, N. E., McVicar, T. R., Vergopolan, N., Berg, A., & Wood, E. F. (2018). Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Scientific Data*, 5(1), 180214. <https://doi.org/10.1038/sdata.2018.214>
- Biau, G., & Scornet, E. (2016). A random forest guided tour. *TEST*, 25(2), 197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Biskaborn, B. K., Smith, S. L., Noetzli, J., Matthes, H., Vieira, G., Streletskiy, D. A., Schoeneich, P., Romanovsky, V. E., Lewkowicz, A. G., Abramov, A., Allard, M., Boike, J., Cable, W. L., Christiansen, H. H., Delaloye, R., Diekmann, B., Drozdov, D., Etzelmüller, B., Grosse, G., ... Lantuit, H. (2019). Permafrost is warming at a global scale. *Nature Communications*, 10(1), 264. <https://doi.org/10.1038/s41467-018-08240-4>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brenning, A., & Trombotto, D. (2006). Logistic regression modeling of rock glacier and glacier distribution: Topographic and climatic controls in the semi-arid Andes. *Geomorphology*, 81(1–2), 141–154. <https://doi.org/10.1016/j.geomorph.2006.04.003>
- Burn, C. R., & Lewkowicz, A. G. (1990). Canadian Landform Examples - 17 Retrogressive Thaw Slumps. *The Canadian Geographer/Le Géographe Canadien*, 34(3), 273–276. <https://doi.org/10.1111/j.1541-0064.1990.tb01092.x>
- Cutler, D. R., Edwards, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random Forests for Classification in Ecology. *Ecology*, 88(11), 2783–2792. <https://doi.org/10.1890/07-0539.1>
- Derksen, C., Burgess, D., Duguay, C., Howell, S., Mudryk, L., Smith, S., Thackeray, C., & Kirchmeier-Young, M. (2019). Changes in snow, ice, and permafrost across Canada. In E. Bush & D. S. Lemmen (Eds.), *Canada's Changing Climate Report* (pp. 194–260). Government of Canada.
- Duk-Rodkin, A. (1999). *Glacial limits map of Yukon Territory*. <https://doi.org/10.4095/210739>
- Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, 37(12), 4302–4315. <https://doi.org/10.1002/joc.5086>
- Fisher, J. P., Estop-Aragonés, C., Thierry, A., Charman, D. J., Wolfe, S. A., Hartley, I. P., Murton, J. B., Williams, M., & Phoenix, G. K. (2016). The influence of vegetation and soil characteristics on active-layer thickness of permafrost soils in boreal forest. *Global Change Biology*, 22(9), 3127–3140. <https://doi.org/10.1111/gcb.13248>
- Fulton, R. J. (1989). *Quaternary Geology of Canada and Greenland*. <https://doi.org/10.4095/127905>
- Harris, S. A., French, H. M., Heginbottom, J. A., Johnston, G. H., Ladanyi, B., Sego, D. C., & van Everdingen, R. O. (1988). *Glossary of Permafrost and Related Ground-Ice Terms*.
- Heginbottom, J. A., Dubreuil, M.-A., & Harker, P. (1995). *Permafrost, Atlas of Canada, 5th Edition*. National Resources Canada. <https://open.canada.ca/data/en/dataset/d1e2048b-cdff-5852-aaa5-b861bd55c367>
- Hijmans, R. J. (2022). *terra: Spatial Data Analysis* (R package version 1.6-17). <https://CRAN.R-project.org/package=terra>
- Hjort, J., Karjalainen, O., Aalto, J., Westermann, S., Romanovsky, V. E., Nelson, F. E., Etzelmüller, B., & Luoto, M. (2018). Degrading permafrost puts Arctic infrastructure at risk by mid-century. *Nature Communications*, 9(1), 5147. <https://doi.org/10.1038/s41467-018-07557-4>
- Jorgenson, M. T., & Osterkamp, T. E. (2005). Response of boreal ecosystems to varying modes of permafrost degradation. *Canadian Journal of Forest Research*, 35(9), 2100–2111. <https://doi.org/10.1139/x05-153>
- Kokelj, S. V., Tunnicliffe, J., Lacelle, D., Lantz, T. C., Chin, K. S., & Fraser, R. (2015). Increased precipitation drives mega slump development and destabilization of ice-rich permafrost terrain, northwestern Canada. *Global and Planetary Change*, 129, 56–68. <https://doi.org/10.1016/j.gloplacha.2015.02.008>
- Kokelj, S. V., Jenkins, R. E., Milburn, D., Burn, C. R., & Snow, N. (2005). The influence of thermokarst disturbance on the water quality of small upland lakes, Mackenzie Delta region, Northwest Territories, Canada. *Permafrost and Periglacial Processes*, 16(4), 343–353. <https://doi.org/10.1002/ppp.536>
- Kokelj, S. V., Kokoszka, J., van der Sluijs, J., Rudy, A. C. A., Tunnicliffe, J., Shakil, S., Tank, S. E., & Zolkos, S. (2021a). Thaw-driven mass wasting couples slopes with downstream systems, and effects propagate through Arctic drainage networks. *The Cryosphere*, 15(7), 3059–3081. <https://doi.org/10.5194/tc-15-3059-2021>
- Kokelj, S. V., Kokoszka, J., van der Sluijs, J., Rudy, A. C. A., Tunnicliffe, J., Shakil, S., Tank, S. E., & Zolkos, S. (2021b). Thaw-driven mass wasting couples slopes with downstream systems, and effects propagate through Arctic drainage networks. *The Cryosphere*, 15(7), 3059–3081. <https://doi.org/10.5194/tc-15-3059-2021>
- Kokelj, S. V., Lacelle, D., Lantz, T. C., Tunnicliffe, J., Malone, L., Clark, I. D., & Chin, K. S. (2013). Thawing of massive ground ice in mega slumps drives increases in stream sediment and solute flux across a range of watershed scales. *Journal of Geophysical Research: Earth Surface*, 118(2), 681–692. <https://doi.org/10.1002/jgrf.20063>
- Lacelle, D., Brooker, A., Fraser, R. H., & Kokelj, S. V. (2015). Distribution and growth of thaw slumps in the Richardson Mountains–Peel Plateau region, northwestern Canada. *Geomorphology*, 235, 40–51. <https://doi.org/10.1016/j.geomorph.2015.01.024>
- Lewkowicz, A. G., & Harris, C. (2005a). Frequency and magnitude of active-layer detachment failures in discontinuous and continuous permafrost, northern Canada. *Permafrost and Periglacial Processes*, 16(1), 115–130. <https://doi.org/10.1002/ppp.522>

- Lewkowicz, A. G., & Harris, C. (2005b). Morphology and geotechnique of active-layer detachment failures in discontinuous and continuous permafrost, northern Canada. *Geomorphology*, 69(1–4), 275–297. <https://doi.org/10.1016/j.geomorph.2005.01.011>
- Loh, W. (2011). Classification and regression trees. *WIREs Data Mining and Knowledge Discovery*, 1(1), 14–23. <https://doi.org/10.1002/widm.8>
- McClymont, A. F., Hayashi, M., Bentley, L. R., & Christensen, B. S. (2013). Geophysical imaging and thermal modeling of subsurface morphology and thaw evolution of discontinuous permafrost. *Journal of Geophysical Research: Earth Surface*, 118(3), 1826–1837. <https://doi.org/10.1002/jgrf.20114>
- Natural Resources Canada. (2017). *Lakes, Rivers and Glaciers in Canada - CanVec Series - Hydrographic Features*. Government of Canada.
- NWT Centre for Geomatics. (2019). *Fire History*. NWT Centre for Geomatics.
- O'Neill, H. B., Wolfe, S. A., & Duchesne, C. (2019). New ground ice maps for Canada using a paleogeographic modelling approach. *The Cryosphere*, 13(3), 753–773. <https://doi.org/10.5194/tc-13-753-2019>
- O'Neill, H. B., Wolfe, S. A., & Duchesne, C. (2022). *Ground ice map of Canada*. <https://doi.org/10.4095/330294>
- Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal*, 10(1), 439–446. <https://doi.org/10.32614/RJ-2018-009>
- Porter, C., & et al. (2022). ArcticDEM – Strips, Version 4.1. In *Harvard Dataverse*, V1.
- Prasad, A. M., Iverson, L. R., & Liaw, A. (2006). Newer Classification and Regression Tree Techniques: Bagging and Random Forests for Ecological Prediction. *Ecosystems*, 9(2), 181–199. <https://doi.org/10.1007/s10021-005-0054-1>
- R Core Team. (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Rudy, A. C. A., Morse, P. D., Kokelj, S. V., Sladen, W. E., & Smith, S. L. (2019). A New Protocol to Map Permafrost Geomorphic Features and Advance Thaw-Susceptibility Modelling. *Cold Regions Engineering* 2019, 661–669. <https://doi.org/10.1061/9780784482599.076>
- Schuur, E. A. G., & Mack, M. C. (2018). Ecological Response to Permafrost Thaw and Consequences for Local and Global Ecosystem Services. *Annual Review of Ecology, Evolution, and Systematics*, 49(1), 279–301. <https://doi.org/10.1146/annurev-ecolsys-121415-032349>
- Sladen, W. E., Parker, R. J. H., Kokelj, S. V., & Morse, P. D. (2021). *Geomorphologic feature mapping methodology developed for the Dempster Highway and Inuvik to Tuktoyaktuk Highway corridors*. <https://doi.org/10.4095/328181>
- Sladen, W. E., Parker, R. J. H., Morse, P. D., Kokelj, S. V., & Smith, S. L. (2022). *Geomorphologic feature inventory along the Dempster and Inuvik to Tuktoyaktuk highway corridor, Yukon and Northwest Territories*. <https://doi.org/10.4095/329969>
- Smith, S. L., Wolfe, S. A., Riseborough, D. W., & Nixon, F. M. (2009). Active-layer characteristics and summer climatic indices, Mackenzie Valley, Northwest Territories, Canada. *Permafrost and Periglacial Processes*, 20(2), 201–220. <https://doi.org/10.1002/ppp.651>
- Stevens, C. (2020). *Stability of Critical Infrastructure in Permafrost Environments: A Proactive Approach*.
- Strobl, C., Boulesteix, A.-L., Zeileis, A., & Hothorn, T. (2007). Bias in random forest variable importance measures: Illustrations, sources and a solution. *BMC Bioinformatics*, 8(1), 25. <https://doi.org/10.1186/1471-2105-8-25>
- Sturm, M., Holmgren, J., McFadden, J. P., Liston, G. E., Chapin, F. S., & Racine, C. H. (2001). Snow–Shrub Interactions in Arctic Tundra: A Hypothesis with Climatic Implications. *Journal of Climate*, 14(3), 336–344. [https://doi.org/10.1175/1520-0442\(2001\)014<0336:SSIIAT>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<0336:SSIIAT>2.0.CO;2)
- Wang, J. A., Sulla-Menashe, D., Woodcock, C. E., Sonnentag, O., Keeling, R. F., & Friedl, M. A. (2019). *ABoVE: Landsat-derived Annual Dominant Land Cover Across ABoVE Core Domain, 1984-2014*. ORNL DAAC.
- Ward Jones, M. K., Pollard, W. H., & Jones, B. M. (2019). Rapid initialization of retrogressive thaw slumps in the Canadian high Arctic and their response to climate and terrain factors. *Environmental Research Letters*, 14(5), 055006. <https://doi.org/10.1088/1748-9326/ab12fd>
- Wies, C., Miltenberger, R., Grieser, G., & Jahn-Eimermacher, A. (2023). Exploring the variable importance in random forests under correlations: a general concept applied to donor organ quality in post-transplant survival. *BMC Medical Research Methodology*, 23(1), 209. <https://doi.org/10.1186/s12874-023-02023-2>
- Wildland Fire Management - Government of Yukon. (2014). *Fire History*. Geomatics Yukon.
- Young, J. M., Alvarez, A., van der Sluijs, J., Kokelj, S. V., Rudy, A., McPhee, A., Stoker, B. J., Margold, M., & Froese, D. (2022). Recent Intensification (2004–2020) of Permafrost Mass-Wasting in the Central Mackenzie Valley Foothills Is a Legacy of Past Forest Fire Disturbances. *Geophysical Research Letters*, 49(24). <https://doi.org/10.1029/2022GL100559>