USING EXPLAINABLE AI (XAI) TO ASSESS THE POTENTIAL OF SATELLITE REMOTE SENSING FOR ESTIMATING ACTIVE LAYER THICKNESS (ALT)

Michael Merchant[1], Lindsay McBlane[1]
[1]Ducks Unlimited Canada, 10525 170 Street, Suite 300, Edmonton, Alberta, Canada, T5P 4W2

ABSTRACT

Permafrost environments are increasingly threatened by warming temperatures. Active Layer Thickness (ALT) is a critical variable because it influences several environmental processes, particularly carbon storage and cycling. Remote Sensing (RS) techniques hold great potential for mapping ALT over remote, challenging, and inaccessible permafrost regions, especially in comparison to in situ field surveys. In this study, various RS sources (radar, optical, topographic) were assessed for ALT estimation. ALT samples were obtained from airborne radar imaging (P- and L-Band) through the NASA Arctic Boreal Vulnerability Experiment (ABoVE), and then upscaled to satellite observations using machine learning (ML) modelling. Using Explainable AI (XAI) methods, we provide interpretation of our ML modeling by assessing the relevance of 116 input RS features. The Shapley additive explanations (SHAP) method was utilized for this interpretation task. Temperature and topography features provided the greatest explanation in ALT variability. Therefore, this study provides insight into the potential of satellite RS for estimating ALT, and the relative importance of RS features for this challenging task.

Index Terms— Arctic, Artificial Intelligence, Explainable AI; Machine Learning, Permafrost, Remote Sensing

1. INTRODUCTION

Permafrost, which is defined as ground remaining at 0°C for two or more consecutive years, underlies an estimated 24% of the northern hemisphere [1]. At these high latitudes, the maximum thaw depth of soils above the permafrost layer, known as Active Layer Thickness (ALT), is an essential variable since it is strongly related to changes in permafrost conditions. As permafrost degrades, the active layer expands into its upper region. Vegetation, topography, hydrology, soil, and climatological parameters act as important controls on ALT [2]. In particular, air temperatures exert large influences on ALT, both inter- and intra-annually. Warmer temperatures increase the duration and depth of thawing, hence why climate change is of concern for ALT conditions. Changes to ALT may have profound impacts on the release of carbon stored in permafrost soils [3]

Quantifying ALT over large areas is challenging since it is a below ground variable, above which exists a mosaic of heterogeneous hydro-ecological conditions (e.g., vegetation, moisture, etc.). Local scale in situ techniques, such as ground penetrating radar (GPR), have been proven effective for ALT mapping [4]. However, these methods, in addition to human-based sampling (e.g., point-based field measurements from probing), are labor intensive, expensive, and thus not reasonably scalable over large spatial extents. On the other hand, remote sensing (RS) measurements, such as from Earth observation (EO) satellites, offer a desirable alternative that is scalable and repeatable. A number of studies have demonstrated the capabilities of RS for local ALT estimation [5], but few have done so over large, regional extents.

In this study, we explore the potential of satellite-based RS for estimating ALT over a high latitude region comprised of both the arctic and boreal biomes, and consisting of various permafrost terrain types. The goal of this study was to identify a potential path forward to mapping ALT over large spatial extents, by identifying the most important RS variables related to ALT. To do this, we employed both machine learning (ML) modelling and explainable artificial intelligence (XAI) techniques. The latter can effectively provide insights, interoperability, and transparency of ML model decision making processes [6].

2. MATERIALS AND METHODS

2.1. Study area

This study was conducted over mainland Northwest Territories (NWT), Canada. This territory falls within three permafrost zones: continuous, discontinuous, and sporadic. Increasing air temperatures, along with changes to precipitation, are driving permafrost thaw in this region [7].

2.2. Airborne ALT measurements

Measurements of ALT used for ML model development and evaluation were obtained from airborne RS, by Chen et al. [8] through the NASA Arctic Boreal Vulnerability Experiment (ABoVE [9]). ALT profiles were retrieved from simultaneous L- and P-band synthetic aperture radar (SAR) instruments onboard NASA's Uninhabited Aerial Vehicle Synthetic Aperture Radar (UAVSAR) system [10]. Research has shown that these UAVSAR ALT estimates compare well with in situ GPR measurements [11]. 18 sites (i.e., flight paths) across the NWT, and their corresponding

ALT retrievals, were used in this study (Figure 1). These profiles reflect the year 2017.

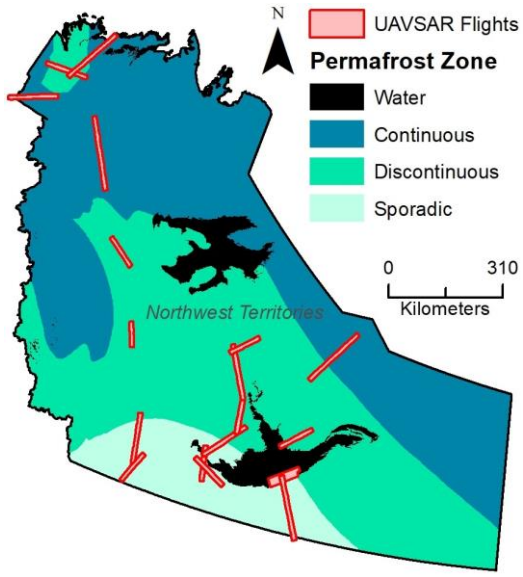

Figure 1: Study area, mainland Northwest Territories, Canada. UAVSAR airborne flight paths are shown.

The UAVSAR measured ALT profiles (i.e., rasters) ranged from 0 to 150 cm depths across the NWT. To obtain ALT samples, we used a random stratification sampling approach. 2500 randomly generated sampling points were created within each 10 cm depth bin (e.g., 0 to 10 cm bin, 10 to 20 cm bin, etc.). This resulted in 37500 randomly generated sampling points. For each point, ALT was drawn from the intersecting UAVSAR ALT raster. These ALT samples were used for ML model tuning and interpretation.

2.3. Satellite datasets

Several satellite-based RS datasets were assessed in this study for upscaling the airborne ALT measurements. Satellite datasets included C-band SAR from Sentinel-1, L-Band SAR from ALOS-PALSAR, and electro-optical data from Landsat-8. These RS sources were chosen since SAR is sensitive to hydrology and can penetrate soil surfaces [12], and Landsat-8 provides reflectance and thermal measurements correlated to sub-surface ground conditions [13]. For each RS data source, a time-series of images was collected using Google Earth Engine (GEE [14]), and then statistical descriptors (mean, median, minimum, maximum, standard deviation, coefficient of variation, and variance) were derived for each band/mathematical indice processed (Figure 2). Statistical descriptors included mean, median, minimum, maximum, standard deviation, variance, and coefficient of variation. Topographic data from the ArcticDEM was also included in our analysis. In total, 116 RS input variables were considered as predictors for ALT estimation (Table 1).

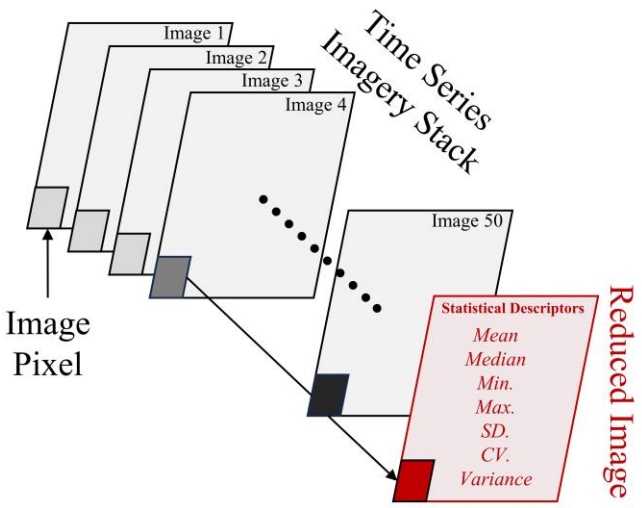

Figure 2: Statistical descriptors computed from RS time-series image stacks.

Table 1: RS variables used for ALT prediction modelling.

RS Source	Variables	Statistical Descriptors
Sentinel-1	σ°VV, σ°VH, VV/VH	Mean, median, min., max., SD, CV, variance
ALOS PALSAR	σ°HH, σ°HV, HH/HV	
Landsat-8	Blue, Green, Red, NIR, SWIR1, SWIR2, Surface Temp, Albedo, NDVI, NDWI	
ArcticDEM	Elevation, Slope, Aspect, TPI	-

2.4. Machine learning modeling

The Random Forest (RF) ML algorithm was chosen for ALT modelling, since it has proven to be robust, accurate, and can handle non-linear relationships and multi-dimensional RS data [15]. A RF model was built using Python programming and the scikit-learn package. Important hyperparameters were tuned using GridSearchCV, which is a technique that searches all possible parameter combinations based on *k* iterations (i.e., subsets). We used the *k*-fold cross validation method for this, with *k*=5.

2.5. Model interpretation using XAI

XAI methods were used for interpreting the initial results of our RF model, since these ML models are often considered "black box". We used the Shapley Additive Explanations (SHAP) algorithm from the shap Python package for visualizing this interpretation. Borrowed from cooperative game theory, Shapley values assess the additive contribution of each input RS variables to a model's prediction [16]. An advantage of SHAP is its ability to interpret both global and local features [17]. Thus, SHAP can reveal the importance of features, and the importance rankings and contributions can be obtained. This holds value as a potential feature selection method for ALT modelling.

3. RESULTS

3.1 Model optimization and Cross-Validation

The hyperparameter tuning process identified four optimized parameters (i.e., settings) for the RF model (Table 2). These parameters were chosen based on the highest achieved correlation coefficient (R^2) from the k-fold cross-validation process, which was R^2 0.432.

Table 2: Hyperparameter values assessed for RF modelling and their final optimized value.

Hyperparameter	Values assessed	Optimal value
'n_estimators'	[25, 50, 75, 100]	100
'max_depth'	[None, 10, 20, 30]	None
'min_samples_split'	[2, 5, 10, 15]	2
'min_samples_leaf'	[1, 2, 4, 8]	2

3.2 Model Interpretation

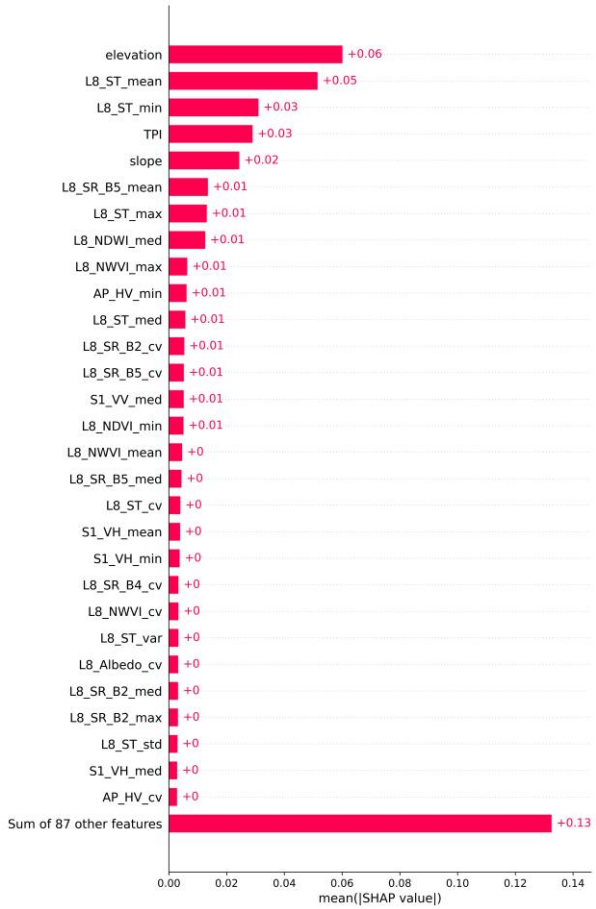

Figure 3: Global feature importance plot, showing the mean absolute SHAP value for the top 30 RS variables.

The SHAP global feature importance plot is presented in Figure 3, which represents the mean Shapley values per input RS feature (i.e., variable), calculated over all samples. This also represents the average impact on ML model output magnitude. Features are plotted by decreasing importance, since the larger the SHAP value, the more important the variable is. Figure 3 indicates that elevation, Landsat-8 mean and minimum surface temperature (ST), Topographic Position Index (TPI), and slope, are the five most important features for predicting (i.e., explaining) ALT variability.

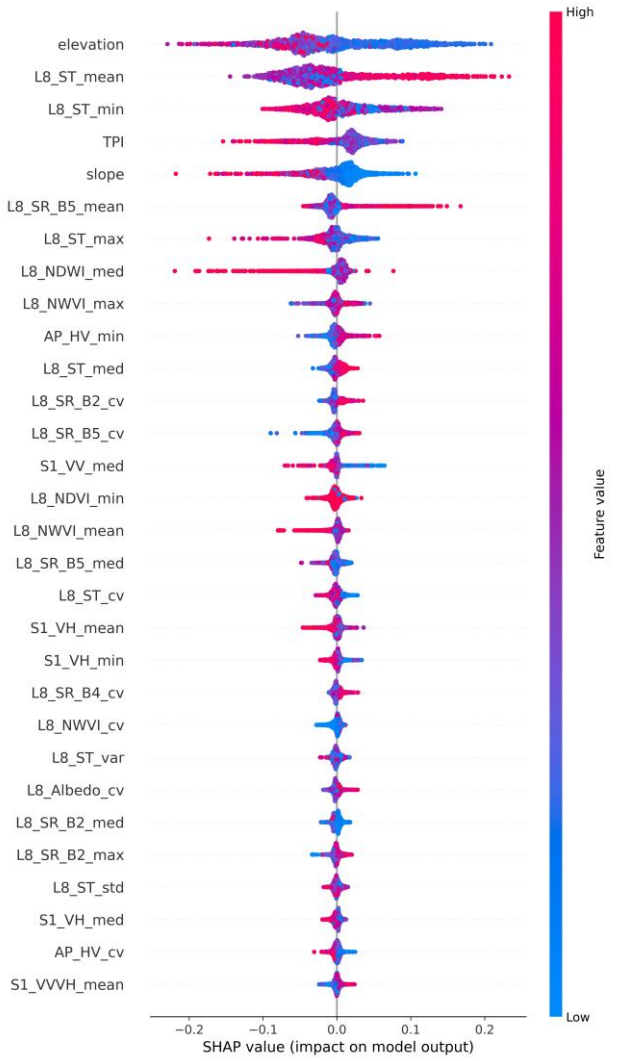

Figure 4: Local feature importance, shown using a SHAP beeswarm summary plot. Positive SHAP values are indicative of deeper ALT, with negative values indicated of shallow ALT.

While average SHAP feature importance is fast and easily interpretable, it does not provide information beyond global importance. Thus, the SHAP summary plot, which combines feature importance with feature effects, is presented in Figure 4. Every point represents the Shapley value of an individual RS variable (i.e., feature) and for a single sample. The SHAP value identifies the probability of different ALT depth predictions, based on the negative or positive correlation to ALT. The most important variables are at the top of the plot. Less important variables, with

SHAP values closer to 0, are at the bottom. This plot shows (1) that for topography, low elevations at lower landscape positions (TPI) and with lower slopes are more likely to result in deeper ALT predictions, and (2) with surface temperatures (ST), high temperatures (both average and minimums, i.e., warmer conditions) are also related to deeper ALT.

4. DISCUSSION

The results of our ML modelling, which was explored using XAI techniques (SHAP), indicated that satellite-collected topographic and temperature RS variables best explain the variability in ALT conditions. The mean SHAP values for the five main explanatory variables, which were all 0.02 or greater (Figure 3) and with wide distributions (Figure 4), provide important clues related to the controls on ALT. Model cross-validation results (R^2 of 432) also suggest that satellite RS measurements collected from multiple sources (radar, optical, and topographic) has the potential to provide relatively good estimates of ALT. This initial validation prediction is encouraging, especially considering that ALT was modeled over a large area and over diverse permafrost conditions (e.g., continuous, discontinuous, and sporadic permafrost zones; Figure 1).

Following this initial work, which focused on ML model evaluation and explainability, we intend on leveraging the valuable insight derived from our XAI methods to design an optimized, efficient, and scalable ML model for ALT mapping with satellite data sets. The SHAP results will be used to select the most important features for this task, whereas CV results will be used to select optimal ML model hyperparameters. Moreover, cloud-based resources provided by GEE will be used to apply these findings over the entire extent of our study area, mainland NWT.

4. CONCLUSIONS

ALT is a critical variable to monitor since it is strongly reactive to climate warming. In situ methods are accurate, but generally not scalable. RS, in contrast, is cost-efficient and can provide wide area coverage. In this study, we examined a large number (116) of RS features, collected by multi-source EO satellite data sets, for their ability to predict ALT via ML modeling. Using novel XAI methods (SHAP), our analysis showed that topography and temperature RS features are critical to estimating ALT. This study provides a promising path forward towards large-scale mapping of ALT over diverse permafrost regions. Cloud-based resources, such as the GEE, can be leveraged for applying these findings over broad spatial scales.

5. REFERENCES

[1] Zhang, T., Heginbottom, J. A., Barry, R. G., & Brown, J. (2000). Further statistics on the distribution of permafrost and ground ice in the Northern Hemisphere. Polar geography, 24(2), 126-131.

[2] Luo, D., Wu, Q., Jin, H., Marchenko, S. S., Lü, L., & Gao, S. (2016). Recent changes in the active layer thickness across the northern hemisphere. Environmental Earth Sciences, 75, 1-15.

[3] Jiang, X., Cai, H., & Yang, X. (2023). Response of active layer thickening to wildfire in the pan-Arctic region: Permafrost type and vegetation type influences. Science of the total environment, 902, 166132.

[4] Sudakova, M., Sadurtdinov, M., Skvortsov, A., Tsarev, A., Malkova, G., Molokitina, N., & Romanovsky, V. (2021). Using ground penetrating radar for permafrost monitoring from 2015–2017 at calm sites in the Pechora river delta. Remote Sensing, 13(16), 3271.

[5] Howard, H. R., Manandhar, S., Wang, Q., Mcmillan, J. M., Qie, G., Liu, X., ... & Wang, G. (2022). Spatially characterizing land surface deformation and permafrost active layer thickness for Donnelly installation of Alaska using DInSAR and MODIS data. Cold Regions Science and Technology, 196, 103510.

[6] Linardatos, P., Papastefanopoulos, V., & Kotsiantis, S. (2020). Explainable ai: A review of machine learning interpretability methods. Entropy, 23(1), 18.

[7] Mamet, S. D., Chun, K. P., Kershaw, G. G., Loranty, M. M., & Peter Kershaw, G. (2017). Recent increases in permafrost thaw rates and areal loss of palsas in the Western Northwest Territories, Canada. Permafrost and periglacial processes, 28(4), 619-633.

[8] Chen, R. H., Michaelides, R. J., Chen, J., Chen, A. C., Clayton, L. K., Bakian-Dogaheh, K., ... & Zhao, Y. (2022). ABoVE: Active Layer Thickness from Airborne L-and P-band SAR, Alaska, 2017, Ver. 3. ORNL DAAC.

[9] Fisher, J. B., Hayes, D. J., Schwalm, C. R., Huntzinger, D. N., Stofferahn, E., Schaefer, K., ... & Zhang, Z. (2018). Missing pieces to modeling the Arctic-Boreal puzzle. Environmental Research Letters, 13(2), 020202.

[10] Rosen, P. A., Hensley, S., Wheeler, K., Sadowy, G., Miller, T., Shaffer, S., ... & Zebker, H. (2007). UAVSAR: New NASA airborne SAR system for research. IEEE Aerospace and Electronic Systems Magazine, 22(11), 21-28.

[11] Parsekian, A. D., Chen, R. H., Michaelides, R. J., Sullivan, T. D., Clayton, L. K., Huang, L., ... & Schaefer, K.

(2021). Validation of permafrost active layer estimates from airborne SAR observations. Remote Sensing, 13(15), 2876.

[12] Merchant, M. A., Obadia, M., Brisco, B., DeVries, B., & Berg, A. (2022). Applying Machine Learning and Time-Series Analysis on Sentinel-1A SAR/InSAR for Characterizing Arctic Tundra Hydro-Ecological Conditions. Remote Sensing, 14(5), 1123.

[13] Kalinicheva, S. V., & Shestakova, A. A. (2021). Using thermal remote sensing in the classification of mountain permafrost landscapes. Journal of Mountain Science, 18(3), 635-645.

[14] Amani, M., Ghorbanian, A., Ahmadi, S. A., Kakooei, M., Moghimi, A., Mirmazloumi, S. M., ... & Brisco, B. (2020). Google earth engine cloud computing platform for remote sensing big data applications: A comprehensive review. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 5326-5350.

[15] Belgiu, M., & Drăguţ, L. (2016). Random forest in remote sensing: A review of applications and future directions. ISPRS journal of photogrammetry and remote sensing, 114, 24-31.

[16] Shapley, L. A Value for n-person Games. In Contribution to the Theory of Games II; Kuhn, H., Tucker, A., Eds.; Princeton University Press: Princeton, NJ, USA, 1953; pp. 307–317.

[17] Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., ... & Lee, S. I. (2020). From local explanations to global understanding with explainable AI for trees. Nature machine intelligence, 2(1), 56-67.