DOI: 10.21917/ijdsml.2025.0176

REVIEW OF MACHINE LEARNING AND DEEP LEARNING TECHNIQUES FOR MAPPING PERMAFROST DISTRIBUTION USING REMOTE SENSING DATA

Anayet Ullah Dar, Muzafar Rasool and Assif Assad
Department of Computer Science and Engineering, Islamic University of Science and Technology, India

Abstract

Permafrost is an important component of cryosphere, crucial in shaping landscapes and regulating ecological processes. A thorough analysis of permafrost modeling techniques is presented in this work, highlighting the application of machine learning(ML) and deep learning(DL) techniques in permafrost conditions. From studies investigating permafrost patterns throughout Swiss Alps to the formation of new permafrost in response to environmental changes in lakes like Zonag Lake, we analyze the effectiveness of ML and DL in capturing the complex dynamics of permafrost evolution. Despite notable achievements, challenges persist, including the need for more comprehensive training data, consideration of local drivers, and the integration of multidisciplinary approaches beyond traditional image processing. For all, the paper highlights the possibility of ML and DL to incorporate exploratory variables by leveraging remote sensing data and climate data, paving the way for enhanced understanding and prediction of permafrost dynamics in critical regions worldwide.

Keywords:

Permafrost Mapping, Remote Sensing Data, Satellite Imagery, Artificial Intelligence

1. INTRODUCTION

Permafrost is characterized by its state of being continuously frozen, where temperatures remain below freezing for a minimum of two years. Permafrost can be found starting just beneath the surface and extending to depths of several hundred meters. Depending on the surrounding environmental conditions, its depth changes considerably. During the summer, ground temperatures above the frozen ground rise above 0°C for a certain period, while below the permafrost base, there are consistently non-cryotic conditions owing to the impact of geothermal heat flux. The layer of the ground above the permafrost that goes through seasonal cycles of freezing and thawing is known as the Active Layer Thickness (ALT). This dynamic layer is essential for understanding permafrost stability and the related ecological and hydrological processes, as it directly affects the moisture and thermal conditions of the ground. For an illustration of the various permafrost Table levels, refer figure1.

Permafrost, which is soil or rock that has been frozen for minimum of two consecutive years, is a key element of our ecosystem. The significance of permafrost is emphasized by its extensive presence, covering approximately a quarter of the bare land in the Northern part of the globe and over 65% of land areas above 60° N latitude [41] [2]. Permafrost plays a pivotal role in landscape formation, hydrological regulation, and biodiversity maintenance [43]. However, permafrost is sensitive to changes in temperature, and its thawing is one of the most significant consequences of climate change. Recent studies have shown that permafrost thawing can have significant impacts on vegetation cover and ecosystem health. For example, thawing permafrost can lead to the formation of thermokarst lakes, which can change the hydrological regime of the surrounding area and alter the distribution of vegetation [43]. Additionally, permafrost thawing can release significant amounts of carbon trapped therein into the atmosphere, which can contribute to global warming [37]. Mountain slope stabilities can be impacted by the thickening of the active layer, which can increase rock fall activity [17] [31] and rock glacier acceleration [21], ultimately boosting sediment flow rates [22]. Monitoring and understanding the impact of permafrost thawing on vegetation cover and ecosystem health is of great importance to the scientific community. However, Traditional Remote sensing and GIS techniques of identifying permafrost distribution and monitoring changes are often time-consuming, costly, and not always able to detect subtle changes.

Machine Learning (ML) is extensively applied to tackle intricate tasks that are challenging for traditional algorithms to accomplish [35]. Deep Learning (DL) algorithms improve these abilities by automatically identifying a large number of hidden features in the data that go unnoticed by humans, making it possible to detect intricate patterns within the data [25]. The ability of both ML/DL to identify intricate patterns within data has significantly broadened their applications in tasks such as segmenting and classifying visual data. Recently, ML techniques are increasingly being used to advancing our knowledge of permafrost and its impact on the ecosystem as they can learn from enormous amounts of data, making them useful for monitoring and tracking changes in permafrost distribution over time. Various studies have employed ML algorithms such as Logistic Regression [24], Random Forests [20], Support Vector Machines [39], and Neural Networks [30] to analyze permafrost data and predict variations in the distribution of permafrost and ALT.

The primary focus of this survey paper is to review research in Geocryology that makes use of ML and DL techniques alongside satellite imagery and remote sensing data for mapping the geographical spread of permafrost.

An overview of the artificial intelligence (AI) techniques employed in the research reviewed within this article is given in section3. Section 4 concludes by summarizing the recent advancements in ML and DL applied to permafrost research, offering insights into the future prospects for studying permafrost. Finally, section 5 presents a summary of key insights, future prospects and takeaways obtained from the reviewed literature.

This systematic review approach will help researchers understand how studies have addressed the challenges and discoveries related to permafrost Dynamics.

2. METHODOLOGY

This survey paper explores the use of ML and DL algorithms to model the distribution of permafrost in the northwest Himalayas. A detailed review of articles published between 2017 and 2023 was conducted. We explored research that

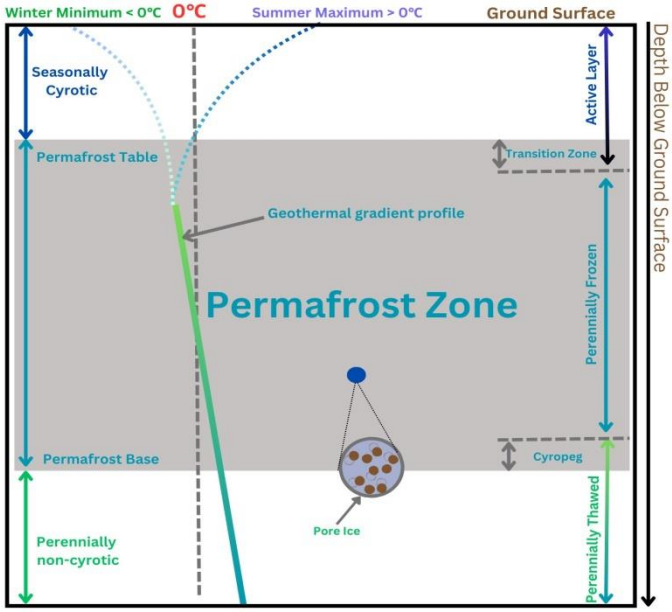

Fig.1. Permafrost Table

addressed the integration of AI in permafrost dynamics, and sourced papers from well-known scientific sources like PubMed, IEEE, and Google Scholar. Original research with notable findings that were published in peer-reviewed journals was a prerequisite for inclusion in our review process. To find relevant studies, we employed specific keywords and phrases such as permafrost distribution, artificial intelligence, deep learning, and permafrost mapping.

A systematic search and screening procedure was used, with full-text evaluations of the shortlisted publications conducted after the titles and abstracts were first evaluated for quality and relevancy. In order to present a coherent narrative on the developments and applications of AI in permafrost dynamics, data from a selection of publications was methodically retrieved, including objectives, techniques, results, and conclusions.

The review begins by providing a concise introduction to various AI algorithms employed in the context of permafrost dynamics. To ensure a comprehensive evaluation, the accuracy achieved by these algorithms in tackling specific research challenges was carefully assessed. This assessment encompassed a thorough examination of the experimental results reported in the reviewed papers, aiming to determine the extent to which the AI and permafrost mapping approaches proved successful in mapping permafrost. Through the implementation of this methodological framework, the goal of this review article is to present a thorough summary of the most recent advancements in the integration of AI and permafrost mapping for environmental applications. The meticulous analysis of the reviewed papers seeks to reveal the potential and limitations of this combined approach, shedding light on the future directions and opportunities for further research and development in this exciting field.

3. ARTIFICIAL INTELLIGENCE

Notable advancements driven by improvement in computing power, processing and refinements of DL/ML models, AI has impacted almost every aspect of existence, including cryospheric sciences. AI is now being used extensively in the field of permafrost monitoring and analysis, climate modeling.

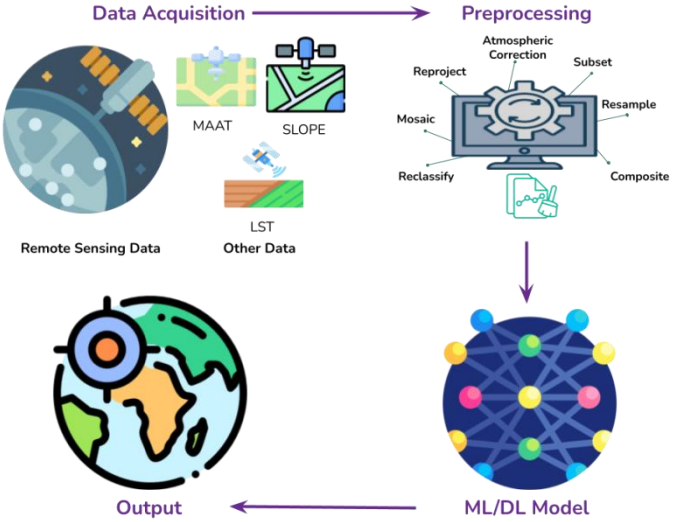

Fig.2. General Workflow Diagram for Permafrost Mapping

The entire process of mapping permafrost employing AI is Summarized in Fig.2, starting with data collection, which involves the collection raw environmental and geospatial data from different sources. Data preparation, which entails cleaning and preparing the data to guarantee its quality and usefulness, comes next. The preprocessed data is then subjected to ML and DL algorithms so as to evaluate and forecast the distribution of permafrost. The algorithm identifies patterns in the data by leveraging its inherent features, enabling it to predict outcomes for new data. The last stage in this process is the creation of a thorough permafrost distribution map, which illustrates the likelihood and spatial extent of permafrost existence based on model predictions. AI algorithms that have been employed in reviewed study are discussed below, as illustrated in the Fig.3.

3.1 MACHINE LEARNING

Although the phrases AI and ML are sometimes used interchangeably, however ML is a subfield of AI that focuses on leveraging pre-existing data and examples to identify patterns and generate predictions for new related challenges [6]. ML Algorithms can be categorized as Supervised, Unsupervised or Semi supervised. This classification is based on whether the algorithms learn from examples where the solutions are provided in the training data or by identifying patterns in the data without any predefined labels [35]. They either come into classification algorithms for classifying data or regression algorithms for making predictions. Some ML algorithms that have been utilized in the reviewed study are discussed hereunder.

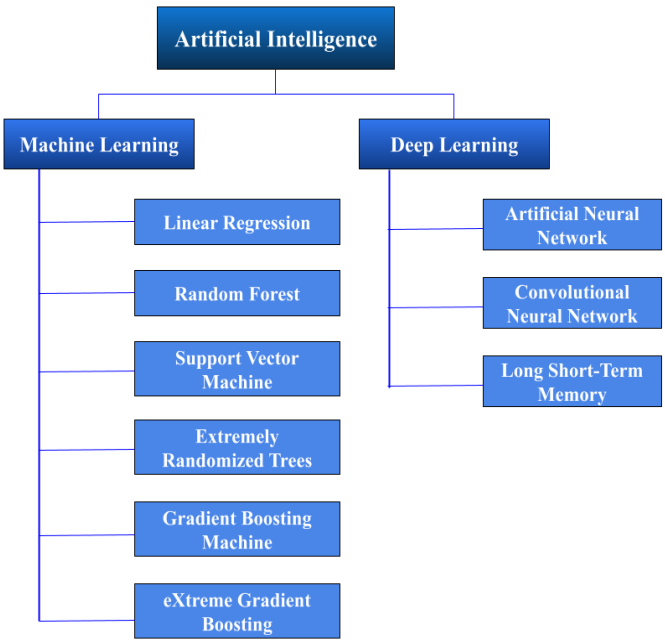

Fig.3. Models Employed in the Permafrost Study

3.1.1 Logistic Regression:

Logistic Regression (LR) is a statistical method for modeling the probability of a binary outcome [24]. LR looks at how different features affect the chance of a specific outcome, performing well when the data exhibits linear separability. The fundamental equation of LR is:

$$P(Y=1|\ X)=\frac{1}{1+e^{-(\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_nx_n)}} \quad (1)$$

where $(Y = 1|X)$ denotes the probability that the dependent variable equals 1 given the vector of independent variables X. The parameters $\beta_0,\beta_1,\ldots,\beta_n$ are estimated from the data. This equation represents the logistic function, which maps any linear combination of the input variables to a value between 0 and 1, thus serving as the basis for predicting binary outcomes [28].

3.1.2 Random Forest:

An ensemble learning technique called Random Forest (RF) uses several decision trees, and the final output is determined by the results of each decision tree based on either majority vote or averaging. [20]. With this approach, a huge number of independent decision trees are built, each with a unique bootstrap sample of the training set. At every split, a random subset of features is chosen for every tree, increasing the variety between the trees. The prediction process involves combining the output from each tree. In the case of classification problems, the final result is decided by majority voting among the individual trees, which can be mathematically expressed as

$$\hat{Y}=\arg\max_c\left(\sum_{b=1}^{B}\mathrm{I}\{T_b(x)=c\}\right) \quad (2)$$

where I is the indicator function, $T_b(x)$ is the prediction of the b^{th} tree, and c stands for the class labels [4]. The result for regression-related problems is the mean of every tree's predictions, which is provided by

$$\hat{Y}=\frac{1}{B}\sum_{b=1}^{B}T_b(x) \quad (3)$$

By combining the advantages of many decision trees, this ensemble method produces high accuracy and tolerance against overfitting, which makes RF a flexible and popular tool for both regression and classification tasks

3.1.3 Support Vector Machine:

Support Vector Machine (SVM) is supervised learning technique primarily intended for problems involving regression and classification [39]. SVM was first developed to deal with linearly separable data; however, it was later improved with the kernel trick to efficiently handle non-linear data. The SVM method seeks to find the best hyperplane that maximizes the margin separating two classes in a high-dimensional feature space [8]. This involves solving the optimization problem shown below:

$$\min_{w,b,\xi}\frac{1}{2}w^{\cdot}\ w+C\sum_{i=1}^{n}\xi_i \quad (4)$$

subject to the constraints:

$$y_i(w^T\phi(x_i)+b)\geq 1-\xi_i,\xi_i\geq 0,i=1,\ldots,n \quad (5)$$

where w denotes the weight vector, b represents the bias term, ξ_i denote slack variables, C represent regularization parameter, y_i denote class labels, x_i are the input vectors, and the function that transform input vectors into a higher-dimensional space is represented by $\phi(x)$. The aim is to balance maximizing the margin and minimizing classification errors.

3.1.4 Extremely Randomized Trees:

Extremely Randomized Trees (ERT), or Extra Trees builds several trees over whole of the dataset during train time [14]. The ERT algorithm introduces strong randomization in both feature and cut-point selection during the tree nodes splitting process. Unlike traditional decision trees, ERT constructs fully randomized trees, not influenced by the output values of the training data. The degree of randomness is adjusted based on specific problem requirements, resulting in adjustments to both bias and variance.

The following is an expression for the ERT splitting criteria:

$$\mathrm{Split}(D)=\arg\max_{j,t}\sum_{i=0}^{n}I(x_{ij}\leq t) \quad (6)$$

where D is the dataset, j is the attribute, t is the threshold, and I is the indicator function.

This process enhances model robustness and generalization, effectively mitigating overfitting, especially in high-dimensional datasets. ERTr's capability to handle large datasets with complex interactions makes it particularly suitable for applications in fields such as remote sensing and geospatial analysis.

3.1.5 Gradient Boosting Machine:

Gradient Boosting Machine (GBM) is an advanced ensemble machine learning technique for regression and classification tasks, that systematically combines several weak learners to create a powerful prediction model. [7]. It operates by continuously tuning new models to reduce the errors of prior models as Gradient descent is used to minimize model loss. Despite its advantages, such as high accuracy and flexibility, GBM can be computationally intensive, prone to overfitting, and sensitive to hyperparameter tuning. Popular implementations include XGBoost, LightGBM, and CatBoost, each offering optimizations for performance and efficiency. The fundamental

equation driving the gradient boosting process for a loss function $L(y,F(x))$ is:

$$F_m(x) = F_{m-1}(x) + \rho_m h_m(x) \quad (7)$$

where $F_m(x)$ is the model at iteration m, $\rho_m$is the step size, and $h_m(x)$ is the base learner added at iteration m. [11] [12]. This approach has shown to be reliable and efficient, particularly when handling raw data and generating models for tasks involving both regression and classification.

3.1.6 eXtreme Gradient Boosting

The Extreme Gradient Boosting (XGBoost) algorithm [7], commonly referred to as XGB, is a powerful machine learning technique designed for classification and regression problems. XGBoost enhances the gradient boosting framework by optimizing the handling of sparse data and implementing a novel tree learning algorithm. XGBoost achieves this through parallel processing, distributed computing, and a regularization technique that reduces overfitting.

In order to minimize the residual errors of the existing models, XGBoost sequentially adds more models. In XGBoost, the prediction function is provided by:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), \quad f_k \in \mathrm{F} \quad (8)$$

where $\hat{y}_i$ is the predicted value, f_k are the individual trees, F represents regression tree space, and K denotes total count of trees. The objective function includes a regularization term to simplify the model, along with a convex loss function to evaluate the gap between the predicted values and the actual targets.

$$L(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \quad (9)$$

where, l represents convex loss function that measures the gap between the predicted value $\hat{y}_i$ and the actual values y_i. The regularization term, given by $\Omega(f) = \gamma^{\cdot} + \frac{1}{2}\lambda \square w\square^2$, aims to limit model complexity, with T being the number of tree leaves and w the vector of predicted values on leaves.

3.2 DEEP LEARNING

A cutting-edge subfield of machine learning known as deep learning has revolutionized both technological and scientific research by taking inspiration from the complex neural networks seen in the human brain [25]. Deep learning algorithms are able to automatically extract complex patterns and representations from data by utilizing large amounts of data and computational power. This paper reviews some of the DL algorithms, which are discussed below.

3.2.1 Artificial Neural Network:

Artificial Neural Networks (ANN) are computational models structured with interconnected nodes known as artificial neurons, arranged in layers: an input layer to receive data, one or more hidden layers where computations occur, and an output layer that produces the network's predictions [34]. The connections between these neurons are weighted, influencing the flow of information from one neuron to other. The neuron's output, denoted as y, is determined by:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \quad (10)$$

where activation function is denoted by f, weights are represented by w_i, input features by x_i, and bias term by b. This equation represents how the input features are processed through the network to generate the final output. This equation represents the transformation of the input features through the network to produce the final output. To reduce the discrepancy between the expected and actual outputs, the network modifies these weights during training in response to the input data [47].

3.2.2 Convolutional Neural Network

Convolutional Neural Networks (CNNs) are specialized DL architectures designed to analyze matrix-based data, like images or sequences. With extra layers like convolutional and pooling layers, this multi-level perceptron assists in identifying notable features in visual input [30]. CNNs are particularly good at extracting abstract elements from visual data that humans might not be able to identify. This characteristic enhances their usefulness in tasks like image identification and classification. Popular CNN models that have made unique contributions to deep learning for visual recognition tasks are ResNet, MobileNet, AlexNet etc.

A CNN usually has several layers, each of which implements a non-linear transformation [23]. Convolution, the core function of a CNN, is defined as follows:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau \quad (11)$$

This operation can be expressed as follows when applied to discrete data, such as digital images:

$$(f * g)(i, j) = \sum_m \sum_n f(m,n)\, g(i-m, j-n) \quad (12)$$

where, g stands for the kernel or filter, and f for the input image. Convolutional layers apply several filters to the input, generating a collection of feature maps. To lessen their dimensionality while keeping the most noticeable features, these feature maps are subsequently run through pooling layers and non-linear activation functions. Optimizing the performance and complexity of the network requires careful consideration of the depth, stride, and zero-padding hyperparameters.

3.2.3 Long Short-Term Memory:

Long Short-Term Memory(LSTM) [15] networks are adept at processing and modeling sequential data having long-term dependencies by selectively retaining and updating information. LSTM is an recurrent neural network (RNN) [45] variant that addresses the issue of vanishing gradient faced by other traditional RNN networks. LSTMs are designed with memory units that maintain information for extended durations. These memory units are equipped with three gates: input gates, output gates, and forget gates, controlling how the information flows through the cell. The equations governing the behavior of these gates and the cell state are as follows:

$$\text{Input Gate: } i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (13)$$

$$\text{Forget Gate: } f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (14)$$

$$\text{Output Gate: } o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (15)$$

$$\text{Cell State Update: } c_t = f_t \square\ c_{t-1} + i_t \square\ \tanh(W_c x_t + U_c h_{t-1} + b_c) \quad (16)$$

$$\text{Hidden State: } h_t = o_t \odot \tanh(c_t) \quad (17)$$

where, the symbol σ denotes the sigmoid activation function, $\odot$ indicates element-wise multiplication, and x_t refers to the input at the t-th time step, hidden state is represented by h_t, cell state by c_t, weight matrices and bias vectors are represented by W, U, and b for respective gates.

4. PERMAFROST MODELLING

The popular Swiss Alps are part of the Alpine range that crosses many European countries, offering abundant biodiversity and playing a vital role in environmental conservation initiatives. Rosablanche, a peak in the Swiss Pennine Alps was selected for a study in order to delineate permafrost extent in the region [9]. ML model have proven useful in analyzing the Swiss Alps' permafrost distribution. A pool of 15 features were considered such as mean annual air temperature (MAAT), Altitude, aspect, slope, lakes, man-made infrastructure etc. some variables showed a strong correlation with the presence or absence of permafrost, while others required for generalization. Out of a total 2,95,680 samples, 6,193 showed permafrost presence while 1,82,173 showed permafrost absence. The remaining 1,07,314 samples were used for testing purposes. SVM model was employed and resulting probabilities were categorized as: less than 0.4 indicated "permafrost absent," 0.4 to 0.7 indicated "permafrost possible," and 0.7 to 1.0 indicated "permafrost probable." To map the possible extent of permafrost, the PERMAL model merged SVM outputs with a rock wall permafrost map [42]. The OAR and AUROC values were determined to be 0.967 and 0.975, respectively, for quantitative analysis. Another study investigated western Swiss Alps with LR, RF, and SVM for mapping mountain permafrost distribution [10]. Aspect, altitude, potential incoming solar radiation (PISR), MAAT, and terrain slope angle were among the environmental factors chosen for permafrost modeling. Evidence of permafrost was also gathered from rock glacier inventorying, and field data, particularly geoelectrical and temperature data, with LR showing a linear relationship with altitude but lacking spatial discontinuity representation. RF exhibited excellent classification performance but differed in smoothness compared to LR, while SVM offered a conservative yet accurate representation with the best reproduction of permafrost discontinuity. AUROC assessments yielded 0.81 for LR, 0.85 for SVM, and 0.88 for RF, indicating robust modeling capabilities. The work emphasized how ML models can effectively understand permafrost distribution and related data properties.

The Qinghai-Tibet Plateau is a high-altitude region in Central Asia. Given Its average elevation of more than 4,500 meters, making it as the world's largest and highest plateau, hence referred to as the "Roof of the World". The QTP is a focus for environmental research because of its susceptibility to climate change, especially in understanding permafrost dynamics and their wider ecological implications. Few investigations focusing on the QTP provide a comprehensive examination of permafrost distribution mapping and associated changes. A study employed ground-penetrating radar (GPR) with 100 and 200 MHz antennas carried a thorough examination of ALT in northeastern QTP [5]. The Study considered mechanical probing, pit observations, and soil thermal property to assessing ALT obtained from GPR. The study revealed that GPR accurately detects ALT with an error margin of ±0.08 meters at sites where midpoints are co-located. The ALT ranged from 0.81 *to* 2.1*m* meters, with an average measurement of 1.32±0.29*m* meters at Eboling Mountain. The average ALT in Yeniu Gou was 2.72 ± 0.88*m* meters. This varied from 1.07*m* meters on slopes facing north to 4.86*m* meters close to the permafrost's lower edge. In peat-covered areas, ALT dropped as elevation increased at rates of –2.1*m/km* in Yeniu Gou and –1.31*m/km* in Eboling Mountain. This rate increased to –4.18*m/km* in Yeniu Gou's mineral soil. In comparison to the slopes facing north, the ALT on the slopes' facing south was thicker, While there was relatively little variation in the peat covered area. Another study investigates the geographical extent of permafrost on the Tibetan Plateau (TP) and the maximum thickness of seasonally frozen ground (MTSFG) and any prospective changes to these features using LR, SVM, and RF [44]. The statistical and ML methods utilized in this study were validated against 106 boreholes across the TP. Results demonstrated high prediction accuracies for permafrost presence, with LR achieving 97.8%, RF 96.7%, and SVM 94.4%, respectively. The study also reveals that during the baseline period, permafrost covers 45.9% of the TP, with projections indicating 25.9% loss of permafrost by 2040s and 43.9% by 2090s. High spatial generalization is demonstrated by SVM, predicting a significant MTSFG decrease in the southwestern TP exceeding 50*cm*by the 2090s. Furthermore, two ML techniques were utilized to simulate the mean annual ground temperature (MAGT) and ALT for the past, present, and future on QTP [29]. These algorithms include the generalized boosting method [32] and RF, the study additionally employed two statistical methods generalized linear modeling and generalized additive modeling. Results indicated reliable simulation of MAGT and ALT, with significant future shrinkage of permafrost under varying climatic conditions. Present permafrost area is estimated at $1.04 * 10^6 km^2$, with 37.3% at risk of disappearance. Under the RCP8.5 scenario, future projections suggest a reduction to 42% of the current area, highlighting pronounced, region specific MAGT and ALT variations. The combined method reveals how permafrost on the QTP responds to climate change. A similar approach utilizing statistical models and ML techniques has also been applied to examine permafrost dynamics on the QTP [38], a region experiencing significant climate-driven transformations. The study investigated permafrost degradation on the TP using empirical models which incluse Temperature at top of permafrost Model (TTOP) [46], Stefan model [33]) and ML techniques (XGB), mapping permafrost distribution and ALT at a resolution of 1 km over four decades. The XGB classification model demonstrates superior performance compared to the TTOP, achieving Accuracy, Precision, and Recall values all exceeding 90.0%, indicating its effectiveness in permafrost distribution estimation. Also, ML models outperforms empirical approaches in modeling permafrost distribution, while the Stefan model better captures ALT. The findings revealed that between 1980 and 2010, the ALT increased by 18.94 cm while the Permafrost area decreased by 15.5%, highlighting spatial patterns of permafrost degradation that reflect greater vulnerability at lower latitudes and altitudes.

Deep learning algorithms have also been used for studying permafrost dynamics. In 2022, the inter-annual variations in ALT was estimated on QTP by RF, CNN, and LSTM [27]. The study

also assessed the seasonal thaw depth during the same period. Meteorological data, on-site ALT measurements, and geospatial data collected between 2002 and 2011 were used as predictive. The Digital Elevation Model, slope, wind velocity, atmospheric pressure, air temperature, soil moisture, rainfall, relative humidity, and the combination of downward short-wave and downward long-wave radiation were used to compute ALT. The Pearson correlation coefficient [19], Spearman correlation coefficient [19] offered directions in selecting the model inputs. The thaw depth reflected stronger positive correlations with temperature, pressure, relative humidity, downward shortwave radiation, downward longwave radiation, and soil moisture when analyzed using Pearson and Spearman correlation coefficients. The study revealed that while RF models perform worse, CNN and LSTM models estimate thaw depth more accurately with increased lagging duration. Also, ALT on the QTP increased between 2003 and 2011, particularly in the north. In 88.7% of the permafrost regions had a deeper thaw in summer, followed by 68.8% in spring, 52.5% in autumn, and 47.5% in winter. While the seasonal thaw depth and active layer thickness in this study were estimated using deep learning algorithms, revealing significant correlations with environmental variables, a separate investigation in 2023 extracted the boundaries of bare permafrost on the TP using random forest models, highlighting the importance of remote sensing in permafrost mapping [26]. By integrating multisource remote sensing data, expert knowledge, and random forest model, the approach achieved high overall accuracies of 90.79% for permafrost classification and 97.47% for bare land extraction. The study revealed that the Perennial permafrost on the TP predominantly occurs above 2000m, with the widest distribution around 5000m, and exhibits elevation-dependent variations influenced by slope direction and solar radiation. Most studies have utilized numerical, equilibrium, and ML models to simulate MAGT in permafrost research. While numerical models offer detailed hydrothermal descriptions, they suffer from high computational demands and uncertainty due to numerous assumptions. Equilibrium models like the TTOP model are widely used but struggle with permafrost boundary measurements. As a result of better permafrost observations, ML models have grown in popularity. They are good at delineating permafrost borders but struggle to represent a wide range of MGT values. To overcome this, a novel approach using the difference between MAAT and MAGT as the target variable has been developed, demonstrating superior accuracy in model validation [48]. Utilizing shared socioeconomic pathways(SSPs) climatic information, the study used ML to forecast the MAGT and ALT of permafrost in the QTP for further assessment. With SSP1-2.6 scenario, projections suggest comparatively stable permafrost whereas for SSP2-4.5, SSP3-7.0, and SSP5-8.5 scenarios, substantial degradation is expected past 2050. Permafrost area losses of 30.15%, 58.96%, and 65.97% are projected for the 2090s compared to the period from 2006 to 2018. Additionally, permafrost in the Qilian range and many Rivers origin area's are identified as particularly vulnerable, with south facing slopes more susceptible to loss, accentuating the slope orientation effect on permafrost presence under climate warming. In a related study [49], the authors examined permafrost formation following the Zonag Lake overflow, which caused the lake's size to shrink by around $100 km^2$. Findings showed rapid permafrost formation on the bare lakebeds, reaching depths of 4.9 m to 5.7 m in the first three years. Simulations suggest continued permafrost development, influenced by ground temperature and thermal amplitude. Unlike bidirectional formation in Arctic taliks, unidirectional permafrost development is observed. Additionally, snow cover and vegetation play significant roles in future permafrost formation on the bare lakebeds.

The Hindu Kush Himalaya (HKH) is another vast and diverse mountain range that contains potential permafrost stretches. The HKH area, sometimes referred to as the "Third Pole" considering its extensive ice and snow reserves, has a significant impact on overall global climatic phenomenon. Despite its significance, the permafrost in this area remains less studied compared to the Arctic and Antarctic regions, presenting a unique set of challenges and opportunities for scientific research. A study visually identified rock glaciers in the HKH region mapping 4,000 samples, each consisting of a square area with a latitudinal extent of 0.05 degrees, corresponding to approximately 5.53 km [36]. Out of the 4,000 samples, 3,432 (86%) received consistent classifications from both mappers: 70% had no rock glaciers, 12% were of poor quality, and 4% contained rock glaciers. This equates to 155 samples, encompassing 702 rock glaciers. Out of these 702 mapped rock glaciers only 5 were lying outside the scope of permafrost zonation index(PZI) [16]. Therefore, the overall analysis indicates a strong alignment between the study and PZI. In Another study, the authors utilized S-2A MSI imagery to identify and map rock glaciers and to analyze the permafrost distribution in Sikkim Himalayas [18]. The study also estimated topographic features and climate factors at rock glacier locations. Correlation analysis revealed that steep slopes and high MAAT contribute to the loss of glaciers at lower altitude, while slope, aspect, and incoming solar radiations influences the mass melting of glaciers at higher altitudes, leading to the formation of glacier-derived rock glaciers. Several LR models were developed, and their results were used to create probability maps depicting the spread of permafrost across the sikkim himalayas. Finally, the degree of likelihood of permafrost ($p \geq 0.5$) within the transect under observation was determined by analyzing the probability maps generated from the LR models. The results revealed that models employed for estimation of permafrost likelihood such as LRM1 predicted 68%($918 km^2$), LRM2 predicted 69%($923 km^2$), LRM3 predicted 52%($705 km^2$), and LRM4 predicted 60%($800 km^2$) of the transect under observation, respectively. This indicates extensive permafrost presence in the Sikkim Himalayas and suggests possible effects of climate change on permafrost in the near future. In a related study [1], the authors developed probability maps of permafrost distribution in the northeastern Himalayan region of Sikkim, utilizing data from remote sensing combined with ML techniques. The study utilized two distinct sets of training data. The first comprised of elevation, slope, and surface reflectance data from Sentinel-2A spectral bands, whereas as the other dataset only contains MAAT and PISR. Eight models, employing LR, RF, SVM, and ANN, with two different input datasets, yielded consistent permafrost extent estimations with varied probability distribution patterns. Despite differences, approximately 60% of the observed area is identified as highly likely to contain permafrost. Additionally, the study also suggests the combined use of surface reflectance, elevation, and slope parameters as another approach for permafrost mapping in Himalayas.

Alaska, with its vast and varied landscapes, is a key region for studying permafrost dynamics due to its extensive discontinuous permafrost zones. Understanding the behavior and spread of permafrost in this specific area is essential for evaluating how climate change affects infrastructure and geomorphic processes. In this context, a study [40] employed ML techniques to create high-resolution maps of permafrost in Alaska's discontinuous permafrost region. Three classifiers, ERTr, SVM, and ANN, were trained using reference data of permafrost presence, which was derived from measurements of electrical resistance and ground temperature. While ERTr exhibits the highest accuracy (70%–90%) for on-site predictions, SVM and ANN models demonstrate better transferability to non-training sites (62%–78%), showcasing the potential of integrating high-resolution spatial data with ML for permafrost mapping.

The Table.1 presents a brief summary of various ML/DL models and the best outcomes they produce in the permafrost analysis and prediction field. The table presents an overview of the several models that have been applied in the field of permafrost analysis, emphasizing the ML/DL models employed, best outcomes achieved, and date of publication. The results make it evident that machine learning methods like XGB, RF, SVM have been essential for achieving high accuracy in analyzing and predicting the permafrost dynamics. Linear regression models, also demonstrate good performance in certain studies despite their simplicity. Notably, research conducted between 2012 and 2024 demonstrate improvement in prediction, with recent ML models such as ERT and GBM demonstrating better regression performance and accuracy. These results indicate the employment of ML/DL models in permafrost research and emphasize how important they are for improving our comprehension and prediction of environmental dynamics specifically permafrost phenomenon.

5. DISCUSSION AND CONCLUSION

Traditional modeling approaches, such as numerical and equilibrium models, have limitations in accurately representing permafrost dynamics, prompting the adoption of ML and DL techniques. AI models demonstrate promising capabilities in simulating permafrost distribution and interannual changes in ALT, offering finer-scale insights compared to conventional methods [29] [38].

Table.1. Summary

Sl. No	Model	Best Results
1	SVM	OAR: 0.967; AUROC: 0.975
2	RF	AUROC: 0.88
3	LR	Accuracy: 89.6%
4		Accuracy: 90.8%
5	ANN	AUROC: 9.20; Accuracy: 96%
6	RF	MAGT: RMSE: 0.53, R: 0.69; ALT: RMSE: 0.85, R: 0.71
7		R^2: 0.87; RMSE: 0.55
8		Accuracy: 90.77%; Kappa: 0.806
9	XGBoost	Accuracy: 96.9%;
		Precision: 92.8%; Recall: 96.9
10	ERT	Accuracy: 91%
11		TFactor: R^2: 0.86, RMSE: 0.41; EFactor: R^2: 0.78, RMSE: 3.12

Thus, the studies highlighted the potential of AI techniques in advancing our understanding of permafrost dynamics. Accurate mapping, prediction, and understanding of permafrost dynamics have been made possible by these advanced algorithms at different regions such as Alaska, the Hindu Kush Himalaya, the Qinghai Tibet Plateau, and the Swiss Alps. According to the reviewed studies, ML and DL models perform better in predicting the spatial distribution of permafrost, especially when they make use of diverse high resolution data and climate data. However, the choice of the optimal band combinations is crucial for mapping applications, and as a result, the input of multispectral bands has a significant influence on the prediction accuracy of the AI models, requiring appropriate band combination adjustment [3].

Although the results show promise, a number of issues still need to be resolved, such as the need for more accurate and high resolution training data, and the consideration of local environmental variables. Moreover, the accuracy of these models can be greatly improved by incorporating microwave data alongside standard high resolution remote sensing datasets, and climatic variables [13]. Most of the permafrost mapping models used in the reviewed study make use of simple machine learning models such as LR, RF, SVM etc. However, using advanced AI models is essential to increase the accuracy of permafrost investigations. Also, Future studies should focus on further fine-tuning the techniques in order to effectively address the effects of climate change on these vital regions and advance our knowledge of permafrost dynamics.

REFERENCES

[1] P. Baral and M.A. Haq, "Spatial Prediction of Permafrost Occurrence in Sikkim Himalayas using Logistic Regression, Random Forests, Support Vector Machines and Neural Networks", *Geomorphology*, Vol. 371, pp. 1-8, 2020.

[2] A. Bartsch, A. Hofler, C. Kroisleitner and A.M. Trofaier, "Land Cover Mapping in Northern High Latitude Permafrost Regions with Satellite Data: Achievements and Remaining Challenges", *Remote Sensing*, Vol. 8, No. 12, pp. 1-27, 2016.

[3] M.A.E. Bhuiyan, C. Witharana, A.K. Liljedahl, B.M. Jones, R. Daanen, H.E. Epstein, K. Kent, C.G. Griffin and A. Agnew, "Understanding the Effects of Optimal Combination of Spectral Bands on Deep Learning Model Predictions: A Case Study based on Permafrost Tundra Landform Mapping using High Resolution Multispectral Satellite Imagery", *Journal of Imaging*, Vol. 6, No. 9, pp. 1-16, 2020.

[4] L. Breiman, "Random Forests", *Machine learning*, Vol. 45, pp. 5-32, 2001.

[5] B. Cao, S. Gruber, T. Zhang, L. Li, X. Peng, K. Wang, L. Zheng, W. Shao and H. Guo, "Spatial Variability of Active Layer Thickness Detected by Ground-Penetrating Radar in the Qilian Mountains, Western China", *Journal of*

Geophysical Research: Earth Surface, Vol. 122, No. 3, pp. 574-591, 2017.

[6] J.G. Carbonell, R.S. Michalski and T.M. Mitchell, "An Overview of Machine Learning", *Machine Learning*, pp. 3-23, 1983.

[7] T. Chen and C. Guestrin, "Xgboost: A Scalable Tree Boosting System", *Proceedings of International Conference on Knowledge Discovery and Data Mining*, pp. 785-794, 2016.

[8] C. Cortes and V. Vapnik, "Support-Vector Networks", *Machine learning*, Vol. 20, pp. 273-297, 1995.

[9] N. Deluigi and C. Lambiel, "PERMAL: A Machine Learning Approach for Alpine Permafrost Distribution Modeling", *Proceedings of Annual Conference of the Swiss Geomorphological Society*, pp. 47-62, 2013.

[10] N. Deluigi, C. Lambiel and M. Kanevski, "Data-Driven Mapping of the Potential Mountain Permafrost Distribution", *Science of the Total Environment*, Vol. 590, pp. 370-380, 2017.

[11] J.H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine", *The Annals of Statistics*, Vol. 29, No. 5, pp. 1189-1232, 2001.

[12] H. Gao, N. Nie, W. Zhang and H. Chen, "Monitoring the Spatial Distribution and Changes in Permafrost with Passive Microwave Remote Sensing", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 170, pp. 142-155, 2020.

[13] P. Geurts, D. Ernst and L. Wehenkel, "Extremely Randomized Trees", *Machine Learning*, Vol. 63, pp. 3-42, 2006.

[14] A. Graves, "Long Short-Term Memory", *Supervised Sequence Labelling with Recurrent Neural Networks*, pp. 37-45, 2012.

[15] S. Gruber, "Derivation and Analysis of a High-Resolution Estimate of Global Permafrost Zonation", *The Cryosphere*, Vol. 6, No. 1, pp. 221-233, 2012.

[16] S. Gruber and W. Haeberli, "Permafrost in Steep Bedrock Slopes and its Temperature-Related Destabilization Following Climate Change", *Journal of Geophysical Research: Earth Surface*, Vol. 112, pp. 1-11, 2007.

[17] M.A. Haq and P. Baral, "Study of Permafrost Distribution in Sikkim Himalayas using Sentinel-2 Satellite Images and Logistic Regression Modelling", *Geomorphology*, Vol. 333, pp. 123-136, 2019.

[18] J. Hauke and T. Kossowski, "Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data", *Quaestiones Geographicae*, Vol. 30, No. 2, pp. 87-93, 2011.

[19] Z. Jin, J. Shang, Q. Zhu, C. Ling, W. Xie and B. Qiang, "RFRSF: Employee Turnover Prediction based on Random Forests and Survival Analysis", *Web Information Systems Engineering*, pp. 503-515, 2020.

[20] A. Kaab, R. Frauenfelder and I. Roer, "On the Response of Rockglacier Creep to Surface Temperature Increase", *Global and Planetary Change*, Vol. 56, No. 1, pp. 172-187, 2007.

[21] F. Kobierska, T. Jonas, J. Magnusson, M. Zappa, M. Bavay, T. Bosshard, F. Paul and S.M. Bernasconi, "Climate Change Effects on Snow Melt and Discharge of a Partly Glacierized Watershed in Central Switzerland (Soiltrec Critical Zone Observatory)", *Applied Geochemistry*, Vol. 26, pp. 60-62, 2011.

[22] A. Krizhevsky, I. Sutskever and G.E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *Communications of the ACM*, Vol. 60, No. 6, pp. 84-90, 2017.

[23] M.P. LaValley, "Logistic Regression", *Circulation*, Vol. 117, No. 18, pp. 2395-2399, 2008.

[24] Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning", *Nature*, Vol. 521, No. 7553, pp. 436-444, 2015.

[25] X. Li, Y. Ji, G. Zhou, L. Zhou, X. Li, X. He and Z. Tian, "A New Method for Bare Permafrost Extraction on the Tibetan Plateau by Integrating Machine Learning and Multi-Source Information", *Remote Sensing*, Vol. 15, No. 22, pp. 1-12, 2023.

[26] Q. Liu, J. Niu, P. Lu, F. Dong, F. Zhou, X. Meng, W. Xu, S. Li and B.X. Hu, "Interannual and Seasonal Variations of Permafrost Thaw Depth on the Qinghai-Tibetan Plateau: A Comparative Study using Long Short-Term Memory, Convolutional Neural Networks and Random Forest", *Science of the Total Environment*, Vol. 838, pp. 1-9, 2022.

[27] M. Maalouf, "Logistic Regression in Data Analysis: An Overview", *International Journal of Data Analysis Techniques and Strategies*, Vol. 3, No. 3, pp. 281-299, 2011.

[28] J. Ni, T. Wu, X. Zhu, G. Hu, D. Zou, X. Wu, R. Li, C. Xie, Y. Qiao and Q. Pang, "Simulation of the Present and Future Projection of Permafrost on the Qinghai-Tibet Plateau with Statistical and Machine Learning Models", *Journal of Geophysical Research: Atmospheres*, Vol. 126, No. 2, pp. 1-20, 2021.

[29] K. O'shea and R. Nash, "An Introduction to Convolutional Neural Networks", *Proceedings of International Conference on Knowledge Discovery and Data Mining*, pp. 1-11, 2015.

[30] L. Ravanel, F. Allignol, P. Deline, S. Gruber and M. Ravello, "Rock Falls in the Mont Blanc Massif in 2007 and 2008", *Landslides*, Vol. 7, pp. 493-501, 2010.

[31] G. Ridgeway, "Generalized Boosted Models: A Guide to the GBM Package", *Update*, Vol. 1, No. 1, pp. 1-12, 2007.

[32] D. Riseborough, N. Shiklomanov, B. Etzelmuller, S. Gruber and S. Marchenko, "Recent Advances in Permafrost Modelling", *Permafrost and Periglacial Processes*, Vol. 19, No. 2, pp. 137-156, 2008.

[33] J.P. Rosa, D.J. Guerra, N.C. Horta, R.M. Martins and N.C. Lourenco, "*Using Artificial Neural Networks for Analog Integrated Circuit Design Automation*", 2020.

[34] H. Sarker, "Machine learning: Algorithms, Real-World Applications and Research Directions", *SN Computer Science*, Vol. 2, No. 3, pp. 1-9, 2021.

[35] M.O. Schmid, P. Baral, S. Gruber, S. Shahi, T. Shrestha, D. Stumm and P. Wester, "Assessment of Permafrost Distribution Maps in the Hindu Kush Himalayan Region using Rock Glaciers Mapped in Google Earth", *The Cryosphere*, Vol. 9, No. 6, pp. 2089-2099, 2015.

[36] E.A. Schuur, A.D. McGuire, C. Schadel, G. Grosse, J.W. Harden, D.J. Hayes, G. Hugelius, C.D. Koven, P. Kuhry and D.M. Lawrence, "Climate Change and the Permafrost Carbon Feedback", *Nature*, Vol. 520, No. 7546, pp. 171-179, 2015.

[37] T. Shen, P. Jiang, Q. Ju, Z. Yu, X. Chen, H. Lin and Y. Zhang, "Changes in Permafrost Spatial Distribution and

Active Layer Thickness from 1980 to 2020 on the Tibet Plateau", *Science of the Total Environment*, Vol. 859, pp. 1-10, 2023.

[38] S. Suthaharan, "*Machine Learning Models and Algorithms for Big Data Classification*", 2016.

[39] E.A. Thaler, S. Uhleman, J.C. Rowland, J. Schwenk, C. Wang, B. Dafflon and K.E. Bennett, "High-Resolution Maps of Near-Surface Permafrost for Three Watersheds on the Seward Peninsula, Alaska Derived from Machine Learning", *Earth and Space Science*, Vol. 10, No. 12, pp. 1-17, 2023.

[40] M. Trofaier, S. Westermann and A. Bartsch, "Progress in Space-Borne Studies of Permafrost for Climate Science: Towards a Multi-ECV Approach", *Remote Sensing of Environment*, Vol. 203, pp. 55-70, 2017.

[41] D. Vonder Muhll, J. Noetzli and I. Roer, "Permos-A Comprehensive Monitoring Network of Mountain Permafrost in the Swiss Alps", *Proceedings of International Conference on Permafrost*, pp. 1869-1874, 2008.

[42] M.A. Walvoord and B.L. Kurylyk, "Hydrologic Impacts of Thawing Permafrost-A Review", *Vadose Zone Journal*, Vol. 15, No. 6, pp. 1-20, 2016.

[43] T. Wang, D. Yang, B. Fang, W. Yang, Y. Qin and Y. Wang, "Data-Driven Mapping of the Spatial Distribution and Potential Changes of Frozen Ground Over the Tibetan Plateau", *Science of the Total Environment*, Vol. 649, pp. 515-525, 2019.

[44] R.J. Williams and D. Zipser, "A Learning Algorithm for Continually Running Fully Recurrent Neural Networks", *Neural Computation*, Vol. 1, No. 2, pp. 270-280, 1989.

[45] J.F. Wright, C. Duchesne and M. Cote, "Regional-Scale Permafrost Mapping using the TTOP Ground Temperature Model", *Proceedings of International Conference on Permafrost*, pp. 1241-1246, 2003.

[46] B. Yegnanarayana, "Artificial Neural Networks for Pattern Recognition", *Sadhana*, Vol. 19, pp. 189-238, 1994.

[47] M. Zhang, R. Li, W. Pei, Y. Zhou, G. Li and S. Yang, "Permafrost Degradation Risk Evaluation in the Qinghai-Tibet Plateau Under Climate Change based on Machine Learning Models", *Journal of Geophysical Research: Atmospheres*, Vol. 129, No. 2, pp. 1-9, 2024.

[48] Y. Zhang, C. Xie, T. Wu, L. Zhao, J. Wu, X. Wu, R. Li, G. Hu, G. Liu and W. Wang, "New Permafrost is Forming on the Exposed Bottom of Zonag Lake on the Qinghai-Tibet Plateau", *Science of the Total Environment*, Vol. 815, pp. 1-7, 2022.