Machine learning for predicting permafrost and active layer temperatures at Toolik Lake, Alaska

Robert Chance[1], Aymane Ahajjam[2], Jaakko Putkonen[1] & Timothy Pasch[3]

[1]*Harold Hamm School of Geology and Geologic Engineering, University of North Dakota, Grand Forks, North Dakota, United States*

[2]*School of Electrical Engineering and Computer Science, University of North Dakota, Grand Forks, North Dakota, United States*

[3]*Department of Communication, University of North Dakota, Grand Forks, North Dakota, United States*

Abstract

Climate change in Arctic regions threatens to significantly degrade permafrost leading to terrain instability, infrastructure damage, altered ecosystems, and the release of greenhouse gasses. Therefore, monitoring permafrost and active layer temperatures is key to understanding where possible permafrost degradation may occur. In areas with high data fidelity, machine learning (ML) algorithms can be utilized for the prediction of subsurface soil temperatures. Here, we present a site-specific model that uses ML in permafrost regions to predict soil temperatures. First, preliminary statistical analyses of the relationships between soil temperatures at several depths and corresponding lag periods of thermal wave propagation are conducted to select the most useful inputs. These inputs are then fed into the prediction model to estimate the corresponding active layer and permafrost temperatures. Various machine learning models, including K-Nearest Neighbors, Random Forest, Gradient Boosting, and Neural Networks, are investigated along with baseline models such as the persistence model. All models were developed using open datasets from Toolik Lake, Alaska. Results highlight the effectiveness of Gradient Boosting (RMSE = 0.13 °C for permafrost) and Random Forest (RMSE = 1.10 °C for active layer) models in predicting soil temperatures. However, these results required data from in situ ground temperature observations.

1 Introduction

Changing environmental conditions, including increasing air temperatures resulting from climate change, are expected to significantly impact permafrost (defined as any in situ geologic materials existing at a temperature of 0 °C or less for two or more years) and active layer (defined as the permafrost soil thermal layer that seasonally breaches 0 °C) temperatures in the 21st century (Melvin et al. 2016; Lawrence et al. 2005; Nitze et al. 2018).The thawing of permafrost (here defined as the annually freezing and thawing surface layer located above the permafrost table) and active layer temperatures will lead to the melting of ground ice. As such, changes in active layer thickness are the best proxy for identifying areas at risk for permafrost degradation-related hazards such as retrogressive thermokarst slumps, thermokarst lakes, and building foundation settlement (Hong et al. 2014; Nelson et al. 2001). The first step towards predicting where these degradational landforms and hazards will occur is the development of ground temperature models accounting for the environmental variability in a changing Arctic landscape.

In all soils, cyclic temperature fluctuations occur as a result of annual variations in incoming solar radiation, among other factors. Together, these factors create a complex interwoven natural system, known as the surface energy budget (Putkonen 1998). Difficulties in measuring the surface energy budget have led researchers to use alternative means to determine ground temperatures since, in most cases, individual energy and mass fluxes at the ground surface at any given location are poorly understood (Douglas et al. 2021; Euskirchen et al. 2016; Ling et al. 2003; Grünberg et al. 2020; Pomeroy et al. 2006; Anisimov et al. 2002; Walker et al. 2003; Loranty and Goetz 2012). Therefore, physics-based models have been developed to estimate and model subsurface temperatures that subvert around complications that arise while modeling ground temperatures using the surface energy budget.

The physics-based one-dimensional heat conduction equation is ubiquitous in periglacial research (Nakano and Brown 1972; Kane et al. 1991; Ling and Zhang 2004) and resolves general ground temperatures with depth well. Despite the success of the heat conduction equation, the reductionist approach given by the model prevents analysis of complex scenarios containing temperature changes, such as those that involve the melting of ground ice (Putkonen 1998). Additionally, physics-based numerical models may fail in situations where the model intends to calculate a value in which the numerical inputs are not directly or precisely known (Chen et al. 2022a). In research where physics-based numerical models fail because of improper inputs, Machine Learning (ML) may provide an alternative as these models have the advantage in that they are not based on the physical processes but rather rely on statistical relationships between various features of the input and output data.

In the realm of non-permafrost soil research, ML has been successfully used to predict soil temperatures at one meter below the ground surface (Bilgili 2010; Zounemat-Kermani 2013; Feng et al. 2019; Alizamir et al. 2020). However, using ML to determine soil temperatures has continued to be challenging due to the physical complexity and non-linearity of the soil-water system that is experiencing freeze-thaw cycles. Hybrid models in which decomposition techniques are used in conjunction with deep learning have

https://doi.org/10.52381/ICOP2024.135.1

shown promise in the prediction of ground temperatures within the active layer on the Tibetan Plateau (Chen et al. 2022b; Liu et al. 2022). However, to date, there is a dearth of published research on ML-modeled prediction of temperatures of the active layer or permafrost and Alaska specifically or the Arctic in general.

For the reasons listed above, there is significant motivation to create an ML-based prediction of soil temperatures leveraging observed temperature data in permafrost regions. Here we present initial steps towards this goal by determining soil temperature with ML algorithms at a site-specific resolution. To achieve this, we investigate in this paper ML-based techniques utilizing more than 17 years of thermal data to predict subsurface soil temperatures at Toolik Lake, Alaska. Accordingly, a preliminary analysis of temperatures at different depths is first conducted to better understand their relationship and identify the best correlated ones. Following this, different supervised conventional ML models are trained to predict the active layer and permafrost temperatures using the selected inputs. Furthermore, model performance is assessed using various metrics and benchmarked with baseline models to provide a clear understanding of their respective capabilities.

2 Study Sites

In this study, we develop site-specific soil temperature prediction models for Toolik Lake, Alaska utilizing data collected by the Natural Resources Conservation Service (NRCS; Staff 2023). This selection is motivated by the availability of long temporal coverage and high-quality data, particularly, time series of air and ground temperatures spanning 18 years at Toolik Lake.

The Toolik Lake borehole is in Northern Alaska, 115 miles south of Prudhoe Bay and north of the Brooks Range at 68.62°N, -149.61°W and at 759 m elevation. This borehole is part of the University of Alaska/National Science Foundation's Toolik Lake soil climate research station. The mean annual air temperature during the data collection period was -7.6 °C. This region experiences annual periods of zero solar insolation due to the high latitude. The monitoring station at Toolik Lake recorded air temperatures at 1.5 m above the ground surface ($\mathrm{T_{air}}$) with soil temperature probes at 0.6 cm (T_0), 8.7 cm (T_8), 16.0 cm (T_{16}), 23.6 cm (T_{23}), 31.2 cm (T_{31}), 38.7 cm (T_{38}), 46.3 cm (T_{46}), 61.6 cm (T_{61}), 76.8 cm (T_{76}), and 97.8 cm (T_{97}) below the ground surface. The vegetation surrounding the borehole is low shrubs.

In March, the active layer temperatures at 76 cm (T_{76}) beneath the ground surface reach their minimum, with a mean temperature of -9.5 °C, while the maximum temperatures occur in August, with a mean temperature of -0.1 °C. Although the mean maximum ground temperature at 76 cm remains < 0 °C during the data collection period, there are instances in July, August, and September where the temperature exceeds the freezing threshold thus signifying that this depth is a transient permafrost depth. There is a noticeable time lag between air temperatures and T_{42}, with the highest air temperatures observed in July, averaging 8.6 °C. This time lag is attributed to the thermal properties of the soil. Similarly, permafrost temperatures at 97 cm (T_{97}) experience their minimum in March, with a mean temperature of -8.9 °C, and their maximum in September, with a mean temperature of -0.6 °C (Figure 1).

We note that the base of the active layer was chosen by finding the nearest ground temperature level measured at a shallower depth than the permafrost table and that contained complete and continuous data. At the considered location (i.e., Toolik lake), this was at 76 cm below the ground surface.

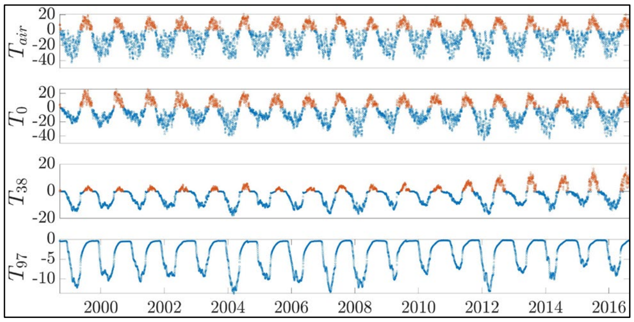

Figure 1. Time series of daily average temperatures at different elevations/depths plotted over 18 years for Toolik lake. Blue dots indicate temperatures < 0 °C and red ones indicate temperatures > 0 °C.

3 Methods

3.1 Soil Temperature Prediction

Let $\{T_d(t)\}_{t=1}^{N}$ denote the daily average soil temperature sequence (of N samples) measured at depth d over N days. Predicting the active layer (or permafrost) temperature T_d^t, at a depth d, and a time index t, can be written as:

$$\hat{\mathrm{T}}_d^t = \mathrm{f}\left(\hat{T}_d^{t-N|t-1}, \epsilon\right) \quad [1]$$

where $\hat{T}_d^t$ is the predicted active layer (or permafrost) temperature at depth d, $T_d^{t-N|t-1} = \{T_d(t-N), \dots, T_d(t-1)\}$ is the sequence of past daily values of the active layer (or permafrost) temperatures at the same depth, f is the prediction model, N is the number of previous days considered as inputs (i.e., lag), and ϵ is the error.

However, we intend in this work to estimate the temperature of the active layer and permafrost without relying on their past values but rather on other predictors, such as ground surface or air temperature values that are more accessible. Hence, Eq. 1 becomes:

$$\hat{T}_d^t = f\left(\mathrm{T}_{d_1}^{t-L_1|t}, \mathrm{T}_{d_2}^{t-L_2|t}, \dots, \epsilon\right), \; d < d_1 < d_2 \quad [2]$$

where $T_{d_1}^{t-L_1|t} = \{T_{d_1}(t-L_1), \dots, T_{d_1}(t)\}$ is the sequence of daily temperatures at depth d_1 over L_1 days, $T_{d_2}^{t-L_2|t} = \{T_{d_2}(t-L_2), \dots, T_{d_2}(t)\}$ is the sequence of daily temperatures at depth d_2 over L_2 days.

Accordingly, three main types of features are employed in this work:

- Air temperature T_{air}^{t}: to provide the average daily air temperature at time index t and acquired at a height of 1.5 m.
- Soil temperatures T_{d}^{t}: to provide average daily soil temperatures at depth d and time index t.
- Time-related features: to provide temporal information reflecting the day of the month D_t ranging from 1 to 31 and month of year M_t ranging from 0 to 12, and year number Y_t. Cyclical changes in the surface energy balance impact the surface energy budget unequally throughout the year. Therefore, using the Julian calendar date for the day on which soil temperatures are to be predicted can bring useful information to the prediction model.

The raw temperature data from Toolik Lake were compiled into a single continuous time series for each location. Once the data were compiled, any temperature data anomalously higher or lower than temperatures within one week before or after the current day were removed from the dataset. Missing data values occurred for no more than one week at their greatest temporal extent, and these values were not estimated from adjacent days to prevent any potential data bias. The time series data are then analyzed to find the first measured depth of the permafrost at each location as well as the approximate active layer thickness. The latter was found by finding the highest mean ground temperature at each depth over the data collection periods. Depths with maximum yearly ground temperatures < 0 °C for a minimum of two years were considered to be permafrost, and active layer thickness was not observed to have increased during the study period, given the vertical resolution of the dataset. Next, time-related features (i.e., day of month, month of year, and year) are read from the corresponding time stamp.

The resulting data are chronologically split into training and testing sets based on an 80/20 ratio and then standardized to have a zero mean and a unit standard deviation. Standardizing the data helps simplify the calculations and improves the model's convergence. Finally, a combination of the preprocessed temperatures (i.e., air temperature or soil temperature at a certain depth), and their corresponding time features, are used as inputs to develop prediction models of the active layer and permafrost temperatures. Different temperature lags are also investigated as inputs to provide direct historical information to the prediction models.

We note that the depth of the active layer was chosen by finding the nearest ground temperature level measured at a shallower depth than the permafrost table and that contained complete and continuous data. At the considered location (i.e., Toolik lake), this was at 76 cm below the ground surface.

3.2 Predictive Models

Two main benchmarks are conducted in this paper to identify the inputs (and corresponding time lags) and models that can produce accurate active layer and permafrost temperature predictions. To accomplish this, the following conventional ML models are considered: Linear regression (LR), K-nearest neighbors (KNN), Support vector regression (SVR), Decision Tree (DT), Random Forest (RF), Gradient Boosting (GB), and Neural Networks (NN). In addition, ARIMA and historical mean persistence models were considered as two baselines for comparison purposes. Accordingly, a conventional ML model is required to outperform both baselines to be considered useful. Numerical modeling of ground temperatures is not used as a benchmark since only ground temperatures are available for Toolik Lake, and more information, such as the thermal conductivity of the soils, is necessary for an educated guess similar to the historical mean and ARIMA models. The mathematical definitions for each of these models are described in detail in (James et al 2022).

3.3 Model Tuning

We conduct a hyperparameter tuning optimization for each of the considered ML models. This optimization aims to find the optimal values for the hyperparameters of each model, which can significantly impact their predictive capabilities and reduce their computational costs. We employ grid search optimization to accomplish this, systematically exploring a predefined set of hyperparameter combinations. A 10-fold cross-validation scheme is used to combat model overfitting and improve performance.

3.4 Performance Evaluation

The prediction performance of all considered models is assessed using four metrics: Root Mean Square Error (RMSE), symmetrical Mean Absolute Percentage Error (sMAPE), R^2, and mean Directional accuracy (DA). Each of these metrics captures and assesses a certain aspect of the prediction models. For instance, RMSE and R^2 are standard regression metrics used extensively in the literature to quantify the performance of a model. More specifically, RMSE quantifies the magnitude of residuals and overall accuracy of the model, while R^2 quantifies how well the model explains the data. Lower RMSE values indicate higher performance, while the opposite is true for R^2. A third metric is sMAPE, which is a variation of MAPE that can bypass its sensitivity to extreme or outlier values and provide a more balanced measure of accuracy. Finally, DA assesses the direction of predictions (upward or downward), rather than their corresponding value.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - f_i)^2} \quad [3]$$

$$sMAPE = \frac{1}{N}\sum_{i=1}^{N}\frac{|y_i - f_i|}{|y_i| + |f_i|} \times 100$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - f_i)^2}{\sum_{i=1}^{N} {y_i}^2} \quad [4]$$

$$DA = \frac{1}{N}\sum_{i=1}^{N} P_i , \quad [5]$$

$$P_i = \begin{cases} 1, if\ (y_{i+1} - y_i) \times (f_{i+1} - y_i) \geq 0 \\ 0, otherwise \end{cases} \quad [6]$$

where y_i is the actual value of a test set instance, i, f_i is the predicted value of the same test set instance, and N is the number of test set instances. Accordingly, the unit of RMSE is °C while sMAPE, R^2 and DA are %.

Now that the main steps to developing the prediction models are described, we proceed to present and discuss the results of our study.

4 RESULTS

4.1 Time Series Analysis

The average Spearman's correlation coefficient computed between different inputs and outputs is shown in Figure 2.

It is clear that current values of soil temperatures are highly correlated with past values of other temperatures from shallower depths (or other heights). More specifically, past values of T_0 seem to be the most correlated with T_{76}, closely followed by those of T_{air}. The correlation coefficients keep increasing for both features until plateauing at around the considered time range limit (i.e., past 60 days) and the value 0.8. The same pattern can be seen for T_{97} with the same features as well as with T_{76}. However, only the recent values of T_{76} are most correlated with the current values of T_{97}.

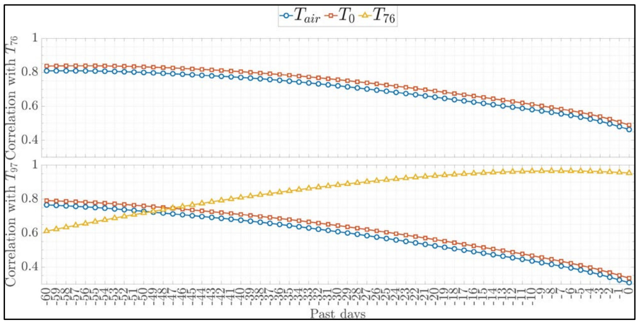

Figure 2. Average Spearman's correlation between different temperature features monitored in Toolik lake, AK. Past values ranging back to 60 days were considered for each feature to compute the average correlation coefficient.

We note that the temporal lags between input features and the soil temperatures chosen for model analysis in the subsequent subsections were guided using Spearman's correlation coefficient values.

4.2 Soil Temperature Prediction

The average prediction performances of all considered models at each soil depth are reported in Tables 1–3. Specifically, Table 1 provides the performance of the baseline models, while Tables 2 and 3 provide the performance of the ML models predicting T_{76} and T_{97}, respectively.

4.2.1 Prediction Using Baseline Models

As can be seen in Table 1, the baseline models had varying prediction performances. In particular, ARIMA consistently achieves the best performances in terms of RMSE, sMAPE, and R^2 across both depths. In addition, it requires only the current temperature of shallower depths (or elevation) to produce such performances. For instance, ARIMA only needed the contemporaneous T_0 to achieve its highest performance of $RMSE = 3.06$°C when predicting T_{76}. In comparison, the historical mean model in the same case achieved its highest performance of $RMSE = 4.01$°C using T_0^{456}, which reflects a lag of more than a year. Nonetheless, the historical mean model can provide predictions that better follow the direction of the targets in all cases.

4.2.2 Prediction Using ML Models

The best-performing baseline models consistently outperform all ML predictors at both soil depths when using only the time-related features (i.e., no information about temperature). This can also be seen when using only the temperature of a much shallower depth to predict the active layer temperature or permafrost temperatures. In both cases, the achieved performances make sense as the employed inputs provide little to no direct information about the target prediction to the prediction models, and hence, can be considered as additional baselines.

Table 1. Baseline models' overall performance (best values are bold). Only the best performing combination of inputs are reported.

Target	Model (inputs)	RMSE	sMAPE (%)	R^2 (%)	DA (%)
T_{76}	Persistence (T_0^{456})	4.01	56.03	25.80	**56.99**
	ARIMA (T_0)	**3.06**	**51.50**	**56.70**	38.37
T_{97}	Persistence (T_{76}^4)	0.44	20.05	99.00	**39.80**
	ARIMA (T_{76})	**0.38**	**17.80**	**99.30**	31.34

Table 2. Average performance of ML models predicting the active layer temperatures at a depth of 76 cm in Toolik lake (best testing performances are bold).

Input	Model	Training		Testing	
		RMSE	R^2(%)	RMSE	R^2(%)
Time	LR	2.80	53.92	2.32	45.25
	KNN	0	100	1.42	79.36
	DT	1.02	93.89	1.71	70.35
	RF	0.97	94.49	1.62	73.10
	GB	0.64	97.55	1.52	76.39
	SVR	2.84	55.59	2.27	57.35
	NN	1.90	80.08	2.10	63.57
T_{air}	LR	3.56	25.87	2.96	10.85
	KNN	3.44	30.55	3.09	2.84
	DT	3.52	27.49	2.99	9.10
	RF	3.51	27.76	2.99	9.40
	GB	3.51	27.88	2.99	9.32
	SVR	3.67	25.90	3.04	23.22
	NN	3.64	27.10	3.08	21.24
T_0	LR	3.45	30.28	2.42	40.51
	KNN	3.33	34.79	2.30	46.18
	DT	3.40	32.21	2.11	54.67
	RF	3.40	32.36	2.13	54.01
	GB	3.40	32.19	2.19	51.13
	SVR	3.52	31.68	3.12	19.52
	NN	3.49	33.05	3.05	22.99
Time+T_0	LR	2.18	72.08	2.23	49.60
	KNN	0	100	1.43	79.18
	DT	1.06	93.42	1.61	73.49
	RF	0.98	94.38	1.55	75.54
	GB	0.64	97.56	1.47	78.11
	SVR	1.96	78.90	1.92	69.55
	NN	3.49	33.05	3.05	22.99
Time+T_0^{15}	LR	1.85	81.07	1.69	76.20
	KNN	0	100	2.14	62.14
	DT	0.46	98.81	1.94	68.84
	RF	0.25	99.66	**1.10**	**89.87**
	GB	0.47	98.78	1.30	85.88
	SVR	1.95	79.14	1.83	72.19
	NN	1.43	88.66	1.80	73.01

Table 3. Average performance of ML models predicting the permafrost layer temperatures at a depth of 97cm in Toolik lake (best testing performances are bold).

Input	Model	Training		Testing	
		RMSE	R^2(%)	RMSE	R^2(%)
Time	LR	2.54	56.36	2.15	52.71
	KNN	0	100	1.57	74.52
	DT	1.36	87.43	1.68	71.08
	RF	1.34	87.74	1.66	71.49
	GB	0.67	96.93	1.54	75.71
	SVR	2.04	71.95	1.89	63.48
	NN	1.61	82.43	1.64	72.37
T_0	LR	3.34	24.30	2.93	11.80
	KNN	3.23	29.11	3.02	6.13
	DT	3.30	26.18	2.96	10.12
	RF	3.30	26.22	2.94	11.53
	GB	3.30	26.38	2.93	11.87
	SVR	3.35	24.15	2.91	13.60
	NN	3.37	23.38	2.82	18.41
T_{76}	LR	0.42	98.79	0.38	98.49
	KNN	0.37	99.03	0.35	98.72
	DT	0.37	99.06	0.36	98.64
	RF	0.36	99.09	0.35	98.70
	GB	0.38	99.00	0.34	98.77
	SVR	0.42	98.78	0.38	98.51
	NN	0.39	98.97	0.34	98.84
Time+T_{76}	LR	0.40	98.88	0.37	98.55
	KNN	0	100	0.41	98.28
	DT	0.11	99.91	0.31	99.00
	RF	0.1	99.93	0.23	99.45
	GB	0.08	99.95	0.30	99.08
	SVR	0.24	99.62	0.32	98.98
	NN	0.24	99.61	0.30	99.07
Time+T_{76}^{30}	LR	0.16	99.81	0.18	99.67
	KNN	0	100	0.37	98.55
	DT	0.04	99.99	0.18	99.66
	RF	0.03	99.99	**0.13**	99.80
	GB	0.01	100	**0.13**	**99.82**
	SVR	0.18	99.78	0.21	99.57
	NN	0.13	99.88	0.20	99.61

However, once the current temperatures of a closer depth (to the target depth) are employed, ML models can perform better than the baselines for both soil depths. Indeed, average test RMSE improvements of 19.03% were achieved across the models when predicting the active layer temperature using the T_0 as input, compared with 5.12% when predicting the permafrost temperature using the active layer temperature as input.

Nonetheless, current temperatures of a closer depth are not enough to accurately predict the active layer or permafrost temperatures. Indeed, even higher improvements are achieved by all considered ML models when provided with a combination of current and past values of the temperatures. When comparing the models, the top prediction performances were consistently achieved using ensemble models in both locations. For instance, the best performance when predicting the active layer temperature is $RMSE = 1.10°\text{C}$ and $R^2 = 89.87\%$ using the input combination $Time + T_0^{15}$ by RF. The best performance when predicting the permafrost temperature is $RMSE = 0.13°\text{C}$ and $R^2 = 99.82\%$ using the input combination $Time + T_{76}^{30}$ with gradient boosting.

5 Discussion

Our results show that there is a higher prediction accuracy when using data features residing closer in physical space to the predicted variable. Under thermally diffusive regimes, it would be expected that ground temperatures at adjacent and shallower depths experience temperature fluctuations before lower depths. This is further evidenced by the temporal lag in shallower ground temperatures that proved to be the best choice of data features when predicting deeper ground temperatures.

While all ML models performed well in predicting soil temperatures when given Julian date and ground temperature time series data in the summer, predictions of ground temperatures in the winter months have reduced accuracy (Figures 3 and 4). One possible cause for this is the presence of snow in the study area that was not captured in the data used in this study. In the winter heat transfer and loss occurs through the emission of radiation from the ground surface, through snow, and then into the air. Since the radiation must pass through snow before the air, and because thermal conductivity of snow differs from that of air, there is a different thermal regime in place during the fall and winter months. As such, future soil temperature prediction using ML models will likely need to account for snow depth along with air temperatures.

Beyond the reduced performance of prediction models during the winter months, a limitation of this study includes the need for ground surface temperature measurements to effectively predict deeper soil temperatures. While air temperatures are routinely measured at airports in the Arctic, very few locations monitor ground surface temperatures. Since the best models presented in this study must use either ground surface or active layer temperatures as inputs rather than air temperatures, the inability to predict ground temperatures solely with the use of air temperatures represents one limiting factor of this study. One method around this would be to predict

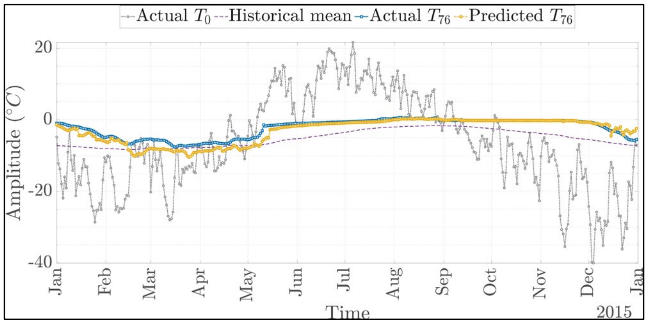

Figure 3. Prediction results of the active layer temperature T_{76} using Random Forest model and Time+T_0^{15}. The historical mean predictions and the actual temperatures over a whole year are plotted for reference.

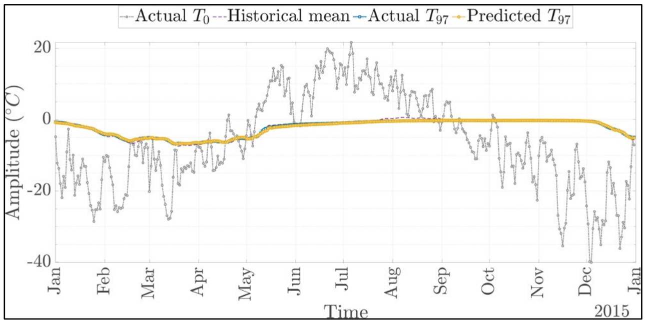

Figure 4. Prediction results of the permafrost temperature T_{97} using Gradient Boosting model and Time+T_{76}^{30}. The historical mean predictions and the actual temperatures over a whole year are plotted for reference.

increasingly deeper ground temperatures based on air temperatures. However, this was not completed for this study, and it is surmised that more than air temperatures and calendar dates would be needed to do this successfully (e.g., snow depth). Another potential method around this limitation may be to utilize an n-factors approach for our input data series (Lunardini, 1978). A second limiting factor to this study is the inability to apply ML models trained at Toolik lake to other locations since the ML models are location-dependent. Future research on the use of ML for location-independent subsurface soil temperature prediction should include additional soil property information such as soil moisture, soil grain size distribution, and surface organic layer depth, amongst other things.

6 Conclusions

Increasing air temperatures due to ongoing anthropogenic climate change are expected to cause the degradation of permafrost in the 21st century. This impending loss of permafrost will lead to the melting of ground ice in regions where it exists today and lead to many associated hazards. To predict where these outcomes may occur, physics-based numerical modeling has been the traditional approach; however, numerical modeling may be insufficient in areas of sparse data, such as the Arctic, where much of the permafrost terrain exists. By making use of existing data series and machine learning methods, we explored the ability of machine learning to estimate the soil temperatures. The predicted temperatures were compared to actual observed soil temperatures to determine which ML methods were best and which variables were needed for optimal performance.

Our research shows that ensemble ML models can successfully predict active layer and permafrost temperatures, capturing the relationships between the air temperature and ground temperatures with the exception for the winter season. For the active layer, soil surface temperatures with a 15 days lag period provided prediction within 1.1 °C of the measured soil temperatures. For permafrost, active layer soil temperatures with lag periods with a 30 days lag period are needed to produce predictions within 0.13°C of the measured soil temperatures. Lastly, these results indicate that if in situ or remote soil temperature observations can be obtained for a given field site, then an AI-based approach may be used to generate active layer and permafrost temperature prediction for those field sites. However, it should be noted that historical soil and air temperatures are required for the training of ML models, and thus these ML models are unable to predict ground temperatures where no historical data is available. Future research should be undertaken to investigate the effectiveness of additional model inputs, such as snow depth and soil properties, for predicting Arctic soil temperatures. Advancing ML models such as those described in this study may further elucidate how Arctic landscapes will evolve with time in a changing Arctic environment.

7 Acknowledgements

DISTRIBUTION A: Approved for Public Release. Distribution is Unlimited. This material is based upon work supported by the Broad Agency Announcement Program and the Cold Regions Research and Engineering Laboratory (ERDC-CRREL) under Contract No. W913E522C0001. This work used advanced cyberinfrastructure resources provided by the University of North Dakota Computational Research Center.

8 References

Alizamir, M., Kisi, O., Ahmed, A.N., Mert, C., Fai, C.M., Kim, N., and El-Shafie, A. 2020. ‘Advanced machine learning model for better prediction accuracy of soil temperature at different depths’, *PLOS ONE* 15(4). doi:10.1371/journal.pone.0231055.

Anisimov, O.A., Shiklomanov, N.I., and Nelson, F.E. 2002. ‘Variability of seasonal thaw depth in permafrost regions: A stochastic modeling approach’, *Ecological Modelling* 153(3), pp. 217–227. doi:10.1016/s0304-3800(02)00016-9.

Bilgili, M. 2010. ‘Prediction of soil temperature using regression and artificial neural network models’, *Meteorology and Atmospheric Physics* 110(1–2), pp. 59–70. doi:10.1007/s00703-010-0104-x.

Chen, T.-C., Penny, S.G., Whitaker, J.S., and Frolov, S.R. 2022a. ‘Correcting systematic and state-dependent errors in the NOAA FV3-GFS using neural networks’, *Journal of Advances in Modeling Earth Systems* [Preprint]. doi:10.1002/essoar.10511972.1.

Chen, L., Liu, X., Zeng, C., He, X., Chen, F., and Zhu, B. 2022b. ‘Temperature prediction of seasonal frozen subgrades based on CEEMDAN-LSTM hybrid model’, *Sensors* 22(15), p. 5742. doi:10.3390/s22155742.

Douglas, T.A., Hiemstra, C.A., Anderson, J.E., Barbato, R.A., Bjella, K.L., Deeb, E.J., Gelvin, A.B., Nelsen, P.E., Newman, S.D., Saari, S.P., et al. 2021. ‘Recent degradation of interior Alaska permafrost mapped with ground surveys, geophysics, deep drilling, and Repeat Airborne Lidar’, *The Cryosphere* 15(8), pp. 3555–3575. doi:10.5194/tc-15-3555-2021.

Euskirchen, E.S., Bennet, A., Bren, A.L, Genet, H., Lindgren, M.A., Kurkowski, T.A., McGuire, A.D., and Rupp, T.S. 2016. ‘Consequences of changes in vegetation and snow cover for climate feedbacks in Alaska and Northwest Canada’, *Environmental Research Letters* 11(10), p. 105003. doi:10.1088/1748-9326/11/10/105003.

Feng, Y., Cui, N., Hao, W., Gao, L., and Gong, D. 2019. ‘Estimation of soil temperature from meteorological data using different machine learning models’, *Geoderma* 338, pp. 67–77. doi:10.1016/j.geoderma.2018.11.044.

Grünberg, I., Wilcox, E.J., Zweiback, S., Marsh, P., and Boike, J. 2020. ‘Linking tundra vegetation, snow, soil temperature, and Permafrost’, *Biogeosciences* 17(16), pp. 4261–4279. doi:10.5194/bg-17-4261-2020.

Hong, E., Perkins, R., and Trainor, S. 2014. ‘Thaw settlement hazard of permafrost related to climate warming in Alaska’, *ARCTIC* 67(1), p. 93. doi:10.14430/arctic4368.

James, G., Witten, D., Hastie, T. and Tibshirani, R.2022. *An introduction to statistical learning: With applications in R*. Springer, Boston, Massachusetts, United States.

Kane, D.L., Hinzman, L.D., and Zarling, J.P. 1991. ‘Thermal response of the active layer to climatic warming in a permafrost environment’, *Cold Regions Science and Technology* 19(2), pp. 111–122. doi:10.1016/0165-232x(91)90002-x.

Lawrence, D.M. and Slater, A.G. 2005. ‘A projection of severe near-surface permafrost degradation during the 21st Century’, *Geophysical Research Letters* 32(24). doi:10.1029/2005gl025080.

Ling, F. and Zhang, T. 2003. ‘Impact of the timing and duration of seasonal snow cover on the active layer and permafrost in the Alaskan Arctic’, *Permafrost and Periglacial Processes* 14(2), pp. 141–150. doi:10.1002/ppp.445.

Ling, F. and Zhang, T. 2004. ‘A numerical model for surface energy balance and thermal regime of the active layer and permafrost containing Unfrozen Water’, *Cold Regions Science and Technology* 38(1), pp. 1–15. doi:10.1016/s0165-232x(03)00057-0.

Liu, Q., Niu, J., Lu, P., Dong, F., Zhou, F., Meng, X., Xu, W., Li, S., and Hu, B.X. 2022. ‘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* 838, p. 155886. doi:10.1016/j.scitotenv.2022.155886.

Loranty, M.M. and Goetz, S.J. 2012. ‘Shrub expansion and climate feedbacks in Arctic Tundra’, *Environmental Research Letters* 7(1), p. 011005. doi:10.1088/1748-9326/7/1/011005.

Lunardini, V.J. 1978. ‘Theory of n-factors and correlations of data’ in *Proceedings of Third International Conference on Permafrost*. National Research Council of Canada, Vol.1, pp. 40–46.

Melvin, A.M., Larsen, P., Boehlert, B., Neumann, J.E., Chinowsky, P., Espinet, X., Martinich, J., Baumann, M.S., Rennels, L., Bothner, A., Nicolsky, D.J., and Marchenko, S.S. 2016. ‘Climate change damages to Alaska Public Infrastructure and the economics of proactive adaptation’, *Proceedings of the National Academy of Sciences* 114(2). doi:10.1073/pnas.1611056113.

Nakano, Y. and Brown, J. 1972. ‘Mathematical modeling and validation of the thermal regimes in tundra soils, Barrow, Alaska’, *Arctic and Alpine Research* 4(1), p. 19. doi:10.2307/1550211.

Nelson, F.E., Anisimov, O.A., and Shiklomanov, N.I. 2001. ‘Subsidence risk from thawing permafrost’, *Nature* 410(6831), pp. 889–890. doi:10.1038/35073746.

Nitze, I., Grosse, G., Jones, B.M., Romanovsky, V.E., and Boike, J. 2018. ‘Remote sensing quantifies widespread abundance of permafrost region disturbances across the Arctic and Subarctic’, *Nature Communications* 9(1). doi:10.1038/s41467-018-07663-3.

Pomeroy, J.W., Bewley, D., Essery, R.L., Hedstrom, N., Link, T., Granger, R.J., Sicart, J.E., Ellis, C., and Janowicz, J. 2006. ‘Shrub tundra snowmelt’, *Hydrological Processes* 20(4), pp. 923–941. doi:10.1002/hyp.6124.

Putkonen, J. 1998. ‘Soil thermal properties and heat transfer processes near Ny-Alesund, northwestern Spitsbergen, Svalbard’, *Polar Research* 17(2), pp. 165–179. doi:10.1111/j.1751-8369.1998.tb00270.x.

Staff, S.S. 2023. *Soil climate research station data for Toolik Lake, Box*. Available at: https://nrcs.app.box.com/s/1vov5vk4mhw3msjtwcsegg9if87rhafu (Accessed: 14 September 2023).

Walker, D.A., Jia, G.J., Epstein, H.E., Raynolds, M.K., Chapin III, F.S., Copass, C., Hinzman, L.D., Knudson, J.A., Maier, H.A., Michaelson, G.J., Nelson, F., Ping, C.L., Romanovsky, V.E., and Shiklomanov, N. 2003. ‘Vegetation-soil-thaw-depth relationships along a low-arctic bioclimate gradient, Alaska: Synthesis of information from the atlas studies’, *Permafrost and Periglacial Processes* 14(2), pp. 103–123. doi:10.1002/ppp.452.

Zounemat-Kermani, M. 2013. ‘Hydrometeorological parameters in prediction of soil temperature by means of artificial neural network: Case study in Wyoming’, *Journal of Hydrologic Engineering* 18(6), pp. 707–718. doi:10.1061/(asce)he.1943-5584.0000666.