A comparative analysis of deep learning models for soil temperature prediction in cold climates

Imanian, Hanifeh; Mohammadian, Abdolmajid; Farhangmehr, Vahid; Payeur, Pierre; Goodarzi, Danial; Hiedra Cobo, Juan; Shirkhani, Hamidreza

Publisher's version / Version de l'éditeur:

https://doi.org/10.1007/s00704-023-04781-x

Theoretical and Applied Climatology, 155, 4, pp. 2571-2587, 2023-12-13

Theoretical and Applied Climatology (2023) 155:2571–2587
https://doi.org/10.1007/s00704-023-04781-x

RESEARCH

A comparative analysis of deep learning models for soil temperature prediction in cold climates

Hanifeh Imanian[1] · Abdolmajid Mohammadian[1] · Vahid Farhangmehr[2] · Pierre Payeur[3] · Danial Goodarzi[1] · Juan Hiedra Cobo[4] · Hamidreza Shirkhani[4]

Received: 4 October 2023 / Accepted: 3 December 2023 / Published online: 13 December 2023

Abstract
Accurate soil temperature prediction in cold climates is crucial for optimizing agricultural practices, hydrological processes, water resource management, minimizing frost damage, and mitigating flood risks. The capacity of deep learning methods to capture intricate patterns and relationships in climate data enhances the accuracy of soil temperature predictions and offers substantial benefits for reducing climate change impacts. In the present study, a comparative analysis of different deep learning techniques, including long short-term memory (LSTM), convolutional neural network (CNN), and multi-layer perceptron (MLP), for predicting the soil temperature is provided. The study examined cold climate areas across Canada, from snowy regions to Arctic conditions. Input datasets were considered both as time series and shuffled order. To comprehensively evaluate the predictive approaches for soil temperature, four machine learning (ML) models—CNN, LSTM, MLP in time series, and MLP on shuffled data—were employed. The results showed ML models using input data as time series have struggled with accurate soil temperature prediction, especially in very cold and polar climates, likely due to the presence of ice layers on the soil, limiting fluctuations near the freezing point. The normalized RMSE (NRMSE) for the CNN, LSTM, and MLP was calculated to be 8.6%. 7.4%, and 6.9%, respectively, and the scatter index (SI) for CNN, LSTM, and MLP was calculated to be 1.0%, 0.9%, and 0.9%, respectively. On the other hand, MLP-shuffled that employs shuffled input data outperformed others with an NRMSE of 5.4% and an SI of 0.7%, by creating a generalized data representation, free from presentation sequence bias. This study showed that predicting soil temperature in very cold climates poses a challenge for machine learning, yet the MLP-shuffled model excels, attaining superior accuracy through the creation of a generalized data representation independent of the sample sequence.

1 Introduction

The accurate prediction of soil temperature plays a crucial role in various fields, such as agriculture, meteorology, environmental science, and land management. This provides valuable insights into plant growth, soil biological activity, hydrological cycles, and ecosystem dynamics (Zeynoddin et al. 2023). By understanding the fluctuations and patterns in soil temperature, researchers and practitioners can make informed decisions regarding irrigation, crop selection, and

✉ Hanifeh Imanian
himania3@uottawa.ca

Abdolmajid Mohammadian
majid.mohammadian@uottawa.ca

Vahid Farhangmehr
vfarhang@uottawa.ca

Pierre Payeur
ppayeur@uottawa.ca

Danial Goodarzi
dgood006@uottawa.ca

Juan Hiedra Cobo
juan.hiedracobo@nrc-cnrc.gc.ca

Hamidreza Shirkhani
hamidreza.shirkhani@nrc-cnrc.gc.ca

[1] Department of Civil Engineering, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[2] Department of Mechanical Engineering, University of Bonab, P.O. Box 55517-61167, Bonab, Iran

[3] School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, ON K1N 6N5, Canada

[4] National Research Council Canada, Ottawa, ON K1A 0R6, Canada

nutrient management. Soil temperature also significantly affects soil technique applications, such as earth-to-air heat exchangers, ground source heat pump systems, and thermal energy storage systems (Yener et al. 2017; Seifi 2021; Abimbola 2021; Ebtehaj et al. 2023). In cold climates, soil temperature prediction plays a pivotal role in optimizing agricultural activities by guiding the optimal timing for planting and harvesting, thereby minimizing frost-induced crop damage. Accurate soil temperature forecasts also contribute to understanding freeze–thaw cycles and their impact on hydrological processes. This information is critical for flood forecasting, water supply management, and mitigating the risks of ice jams and related hazards (Pomeroy et al. 2022). Therefore, soil temperature prediction is a key element in refining water resource management strategies to counteract the potential repercussions of climate change on water availability in cold climates.

Accurate prediction of soil temperature relies on taking into account the various parameters that affect the thermal balance of the soil. Heat transfer between soil and air is a complex process that is influenced by a range of climatic factors, such as air temperature, solar radiation, thermal radiation, wind, evaporation, surface pressure, precipitation, and dew point (Samadianfard 2018; Mehdizadeh 2018).

Traditionally, soil temperature is monitored using ground-based weather stations or manual measurements at specific locations. Although these methods provide localized information, they often lack spatial coverage and temporal resolution, limiting their usefulness for broader applications. In recent years, the advent of remote sensing technologies and the increasing availability of soil temperature datasets have provided new opportunities for predictive modeling over broader data distributions using machine learning techniques (Bonakdari 2019). Historically, machine learning has showcased its potential as a valuable tool for making predictions across diverse areas of hydromechanical engineering (Parveen et al. 2020a, 2020b). As such, machine learning methods that are capable of extracting patterns and relationships from vast amounts of data have revolutionized the field of soil temperature prediction (Hao et al. 2021; Yu et al. 2022).

Taheri et al. (2023) reviewed machine learning techniques implemented for soil temperature estimation, their challenges, and the milestones achieved in this domain. A study by Ebtehaj et al. (2023) presented a novel emotional neural network (ENN) for soil temperature modeling and compared the results with other popular modeling techniques such as generic programming. A study by Belouz and Zereg (2023) investigated the use of an extreme learning machine (ELM) to predict daily soil temperature using only air temperature and relative humidity. Farhangmehr et al. (2023) employed a one-dimensional convolutional neural network (CNN) model to predict the hourly soil temperature under both normal and hot weather conditions. Imanian et al. (2022) delivered a comprehensive and detailed assessment of the performance of a broad spectrum of AI approaches, spanning from well-established methods, such as random forests (RFs) and multi-layer perceptron (MLP), to more advanced deep learning techniques, in soil temperature prediction. Tabrizi et al. (2021) integrated a CNN with an LSTM to predict pavement surface temperature for road management purposes. Hou et al. (2022) employed a combination of two deep learning methods, CNN and LSTM, for hourly temperature prediction. They combined the benefits of CNN to reduce the dimensionality of time-series data, whereas LSTM captured the long-term memory of massive temperature time-series data. Babbar et al. (2023) employed an LSTM model to estimate wheat crops based on the soil and meteorological parameters. Dai et al. (2023) proposed an advanced ensemble deep learning model using an LSTM network to improve the accuracy of road surface temperature prediction. Ju and Liu (2021) applied the attention mechanism to LSTM, enabling it to screen multiple sequences, remove irrelevant redundant information, and capture interaction information between sequences.

Xing et al. (2018) utilized a data-driven model based on a support vector machine (SVM) to predict daily soil temperature in arid, warm, and snowy climates. Delbari et al. (2019) compared SVM and linear regression models for daily soil temperature prediction in different climate conditions, including hyper-arid, arid, semi-arid, Mediterranean, and hyper-humid. Bayatvarkeshi et al. (2021) employed different machine learning models encompassing artificial neural networks (ANN) and co-active adaptive neuro-fuzzy inference systems (CANFIS) in six diverse climate types, including warm summers and mild-to-cold winters.

The above literature review indicates that there is a lack of research on the prediction of soil temperature in cold climates. Existing studies have predominantly employed conventional models rather than advanced machine learning techniques. Overall, the literature gap encompasses the limited utilization and absence of thorough comparisons among deep learning methods for predicting soil temperature in cold climates.

This study aimed to assess the performance of deep learning techniques in predicting soil temperature based on diverse land and atmospheric variables. Three methods of LSTM, CNN, and MLP fed by both time-series and shuffled input data were investigated within the scope of this research. Meanwhile, the study considered areas across cold climates, from snowy regions to Arctic conditions, to address different aspects of the problem.

The remainder of the paper is structured as follows. Section 2 outlines the study areas and the parameters involved, introduces the employed machine learning methodologies, and provides an overview of the study's methodology. This

section also introduces five evaluation metrics that are widely used in the literature to assess the performance of predictive machine-learning models. Section 3 presents the results and subsequently conducts a comparison of them with actual data. This discussion focuses on the performance of various methods using error metrics. Concluding remarks and suggestions for future studies are presented in the final section.

2 Material and methods

2.1 Study area and dataset

The present study focused on investigating soil temperature prediction in cold regions, and for this purpose, three study areas representing distinct climate zones were chosen. The selection was based on the widely recognized Koppen-Geiger climate classification, which provides a robust framework for categorizing climates around the world. The geographical distribution of the study stations is depicted on a country map in Fig. 1.

The first study location was near Ottawa, the capital city of Canada (45.25° N, 75.50° W), located in the southeast of the country, in the province of Ontario. Ottawa, classified as Dfb according to the Koppen-Geiger climate zone classification system, has a humid continental climate featuring cold snowy winters and warm summers [Beck et al. 2018].

Churchill, the second study point, is a coastal city in Canada (58.75° N, 94.00° W), located in the center of the country in the province of Manitoba. Churchill, classified as Dfc, has a subarctic climate featuring very cold, long winters and short, cool summers [Beck et al. 2018].

The last study station was Iqaluit, a city near the Arctic region in Canada (63.75° N, 68.50° W), located in the northeast of the country in the territory of Nunavut. Iqaluit, classified as ET, has a tundra climate, a type of polar climate featuring long, cold winters and brief, cool summers [Beck et al. 2018].

The hourly climate data used in this study were sourced from the freely accessed ERA5 (https://cds.climate.copernicus.eu/, accessed in November 2022), which is a fifth-generation atmospheric reanalysis provided by the Copernicus Climate Change Service at the ECMWF. It is a state-of-the-art reanalysis dataset that combines observations from various sources, including satellites, weather stations, and other measurements, using advanced data assimilation techniques. The data are available in a regular latitude–longitude gridded format with a horizontal resolution of $0.25° \times 0.25°$.

To review the distribution of the soil temperature data at these three stations and their changes over time, considering records from the years 2020 to 2021, a box plot is shown in Fig. 2.

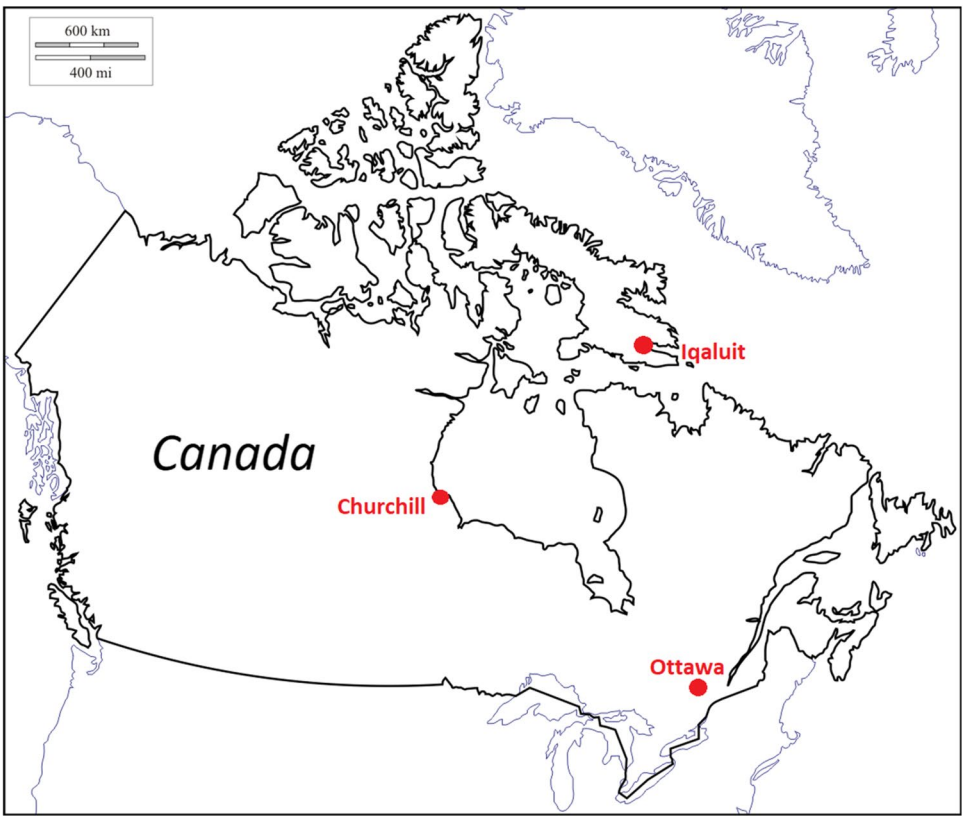

Fig. 1 Geographical distribution of study stations

As shown in Fig. 2, Ottawa usually experiences a temperate climate, with average soil temperatures above freezing. The increase in the average temperature from 2020 to 2021 suggests a warming trend, whereas a decrease in the temperature range might indicate more stable climate conditions. Churchill has a colder climate with under-freezing average soil temperatures. The increase in average temperature from 2020 to 2021 might imply a milder year. The relatively consistent temperature range suggests that extreme climatic conditions were maintained. Ottawa and Churchill both experienced temperature increases from 2020 to 2021, possibly influenced by broader climatic trends. Iqaluit, which is located in a subarctic climate, has an average soil temperature below freezing. The slight increase in the average temperature and the reduced temperature range from 2020 to 2021 indicate a relatively stable cold climate, albeit with some minor fluctuations.

The variables employed in the analysis performed in this study included hourly weather conditions, including air temperature at 2 m above the surface (K), total precipitation (m), surface pressure (Pa), evaporation (m of water), instantaneous wind gusts at 10 m above the surface (m/s), dew point temperature 2 m above the surface (K), surface net solar radiation (J/m^2), and surface net thermal radiation (J/m^2). The data were organized in chronological order and used as the input dataset for each of three study locations of Ottawa, Churchill, and Iqaluit. The out variable would be hourly soil temperature (K) at a depth of 0–7 cm underground. Descriptive statistics of the employed dataset are presented in Table 1.

To show the strength and direction of the relationship between the considered climate variables, the Spearman correlation matrix is illustrated in Fig. 3. The correlation matrix is particularly useful when dealing with nonlinear data, making it a valuable tool in various fields including statistics, data analysis, and machine learning.

The correlation matrix in Fig. 3 suggests that soil temperature (ST) is influenced mainly by air temperature (AT) and dew point temperature (DT), and moderately by solar radiation (SSR) and evaporation (E), which is expected as the ground absorbs heat from the surrounding air, while solar radiation is a direct heat source. Other parameters had weaker effects on soil temperature. This understanding confirms the findings of Xing et al. (2018) that soil temperature is mainly determined by air temperature, solar radiation, and relative humidity.

2.2 Multi-layer perceptron (MLP)

MLPs are artificial neural networks that consist of multiple layers of interconnected neurons. Each neuron receives inputs, performs a mathematical computation, and produces an output, which is then passed to the next layer of neurons. The neurons in an MLP are organized into layers, including one input layer, one or more hidden layers, and one output layer. The hidden layers perform computations on the input data, transforming them into a more abstract representation, whereas the output layer generates the final prediction.

MLPs have some limitations in climate prediction because they struggle to comprehensively manage spatial and temporal dependencies, which are crucial for capturing the complex dynamics of climate systems. MLPs are feedforward networks that do not possess the inherent memory required to handle time-series data effectively.

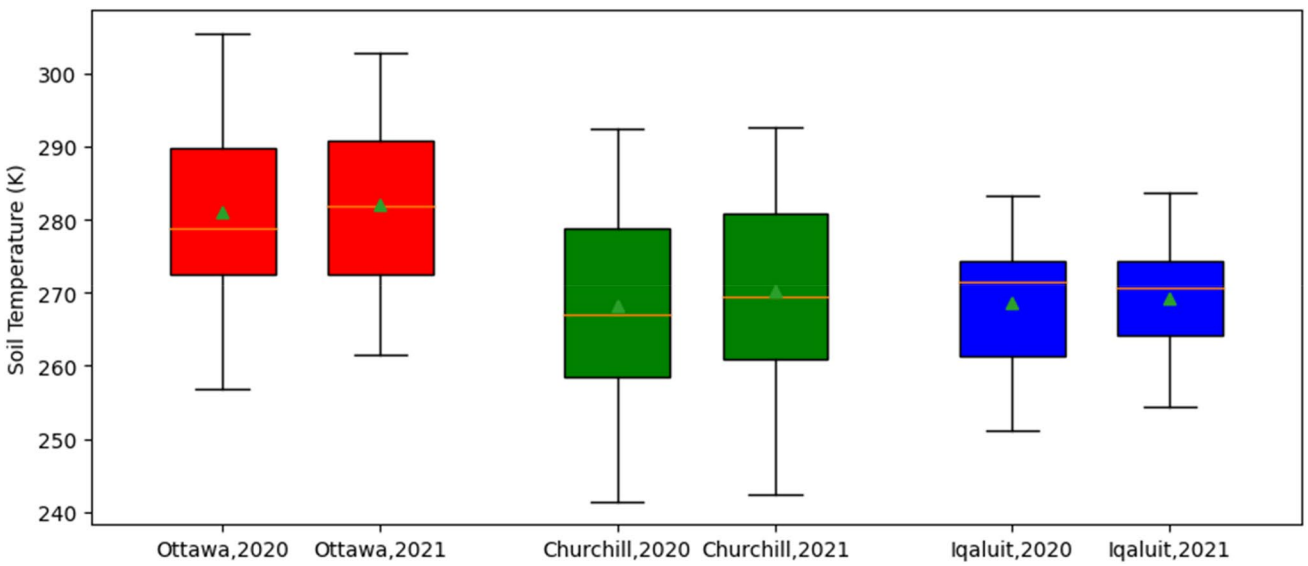

Fig. 2 Box plot of soil temperature distributions in three cities in two consecutive years

Table 1 The descriptive statistics of the employed climate dataset

Station/year	Variable Statistics	Total precipitation (m)	Surface pressure (Pa)	Evaporation (m)	Wind gusts (m/s)	Dewpoint temperature (K)	Solar radiation (J/m^2)	Thermal radiation (J/m^2)	Air temperature (K)	Soil temperature (K)
Ottawa, 2020	Mean	1.2E−04	1.0E+05	−7.3E−05	6.9	275.6	4.6E+05	−1.9E+05	281.1	281.0
	St dev	4.0E−04	8.7E+02	1.0E−04	3.3	11.0	7.0E+05	1.1E+05	11.4	9.9
	Max	5.0E−03	1.0E+05	3.6E−05	20.4	296.9	2.9E+06	5.2E+04	308.9	305.5
	Min	0.0E+00	9.8E+04	−6.1E−04	0.9	240.6	3.6E−12	−4.9E+05	242.9	256.8
	Range	5.0E−03	5.7E+03	6.5E−04	19.5	56.3	2.9E+06	5.5E+05	66.0	48.7
Ottawa, 2021	Mean	1.1E−04	1.0E+05	−7.4E−05	6.5	276.1	4.6E+05	−2.0E+05	281.4	281.9
	St dev	3.9E−04	7.4E+02	1.0E−04	3.1	11.5	7.0E+05	1.1E+05	11.8	9.6
	Max	5.9E−03	1.0E+05	6.7E−05	24.3	297.0	3.0E+06	3.9E+04	305.8	302.9
	Min	0.0E+00	9.7E+04	−5.4E−04	0.5	245.1	3.6E−12	−5.0E+05	248.7	261.5
	Range	5.9E−03	5.3E+03	6.0E−04	23.8	51.8	3.0E+06	5.4E+05	57.2	41.3
Churchill, 2020	Mean	6.0E−05	1.0E+05	−3.1E−05	8.6	264.0	2.4E+05	−1.4E+05	266.9	268.2
	St dev	1.9E−04	1.0E+03	4.4E−05	3.8	13.9	4.4E+05	9.0E+04	13.9	11.7
	Max	3.4E−03	1.1E+05	4.4E−05	29.0	290.2	2.7E+06	8.7E+04	297.3	292.4
	Min	−4.2E+05	9.8E+04	−3.0E−04	0.6	233.1	3.6E−12	−4.2E+05	236.6	241.5
	Range	4.2E+05	7.0E+03	3.5E−04	28.3	57.1	2.7E+06	5.1E+05	60.7	50.9
Churchill, 2021	Mean	6.5E−05	1.0E+05	−3.4E−05	8.5	265.4	2.7E+05	−1.5E+05	268.5	270.3
	St dev	2.2E−04	1.0E+03	4.7E−05	4.0	14.2	4.9E+05	9.4E+04	14.5	11.4
	Max	3.9E−03	1.0E+05	4.7E−05	24.3	293.4	2.7E+06	5.0E+04	298.7	292.6
	Min	0.0E+00	9.8E+04	−3.0E−04	0.8	232.3	3.6E−12	−4.4E+05	236.2	242.4
	Range	3.9E−03	6.6E+03	3.5E−04	23.5	61.1	2.7E+06	4.9E+05	62.5	50.2
Iqaluit, 2020	Ave	6.6E−05	9.9E+04	−2.0E−05	6.0	262.2	2.3E+05	−1.6E+05	265.4	268.7
	St dev	1.9E−04	1.0E+03	2.4E−05	3.1	12.6	4.0E+05	9.1E+04	12.9	7.9
	Max	3.4E−03	1.0E+05	2.6E−05	21.4	283.4	2.5E+06	7.3E+04	293.5	283.2
	Min	0.0E+00	9.5E+04	−1.6E−04	0.8	232.5	3.6E−12	−4.2E+05	236.5	251.2
	Range	3.4E−03	7.9E+03	1.8E−04	20.6	50.9	2.5E+06	4.9E+05	57.0	32.0
Iqaluit, 2021	Ave	6.7E−05	9.9E+04	−1.8E−05	6.2	264.0	2.1E+05	−1.5E+05	267.0	269.3
	St dev	1.6E−04	1.0E+03	2.2E−05	3.1	10.8	3.6E+05	9.7E+04	10.9	6.4
	Max	2.5E−03	1.0E+05	5.2E−05	19.4	282.5	2.5E+06	8.1E+04	293.4	283.8
	Min	0.0E+00	9.5E+04	−1.7E−04	0.7	235.9	3.6E−12	−4.2E+05	240.6	254.3
	Range	2.5E−03	6.6E+03	2.2E−04	18.7	46.6	2.5E+06	5.0E+05	52.8	29.4

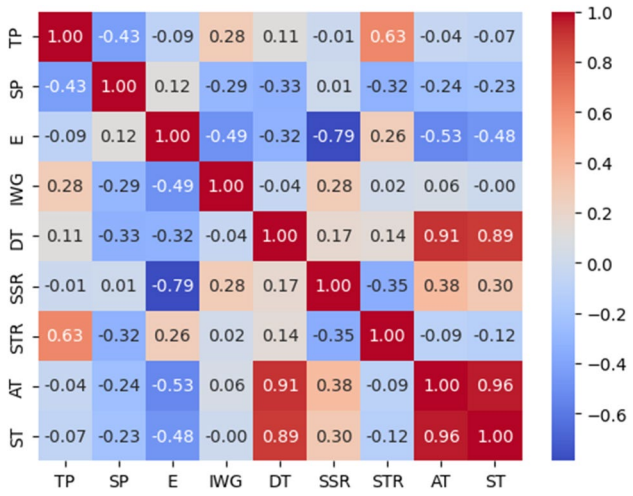

Fig. 3 Correlation matrix of involved climate variables

The network structure of an MLP, including hidden layers, is illustrated in Fig. 4.

The first step in MLP modeling is the design of its architecture. After some trial and error, four dense layers were considered for the present study, including 128, 64, 32, and 1 neuron, respectively. Different activation functions such as rectified linear unit (ReLU), sigmoid, and tanh have been tested, and finally, ReLU has been chosen for the first three layers, while the last one was considered linear. The mathematical formulations of common activation functions in neural networks are provided as

$$\text{ReLU function } f(x) = max(0, x) \tag{1}$$

$$\text{Sigmoid function } f(x) = \frac{1}{1+e^{-x}} \tag{2}$$

$$\text{Tanh function } f(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{3}$$

$$\text{Linear function } f(x) = x \tag{4}$$

Each MLP network has some hyperparameters that should be meticulously tuned to obtain the best performance from the model. The Adam optimization function was considered for its robustness and efficiency after trying different optimizers, such as stochastic gradient descent (SGD) and root mean squared propagation (RMSprop). A range of 0.0001 to 0.1 was attempted for the learning rate of the optimizer, and the best value of 0.001 was found, which led to faster convergence.

Because smaller batch sizes can help with better generalization, the present study tested with 64 and 32, and considered 16 after tuning. However, very small batch sizes may slow the training process. The early stopping technique was used to determine the ideal number of epochs during the training process. When the loss becomes stable or starts to increase, this indicates that the model might have already learned the relevant patterns from the data. Stopping the training at this point prevents the model from further fitting the noise and ensures that it captures the most relevant patterns. On the basis of this strategy, 200 epochs were selected to train the model. When shuffling was applied to the MLPs, the training data were randomly reordered before each epoch during the training process.

In climate prediction, shuffling data may disrupt inherent patterns if the data exhibit temporal dependencies or seasonality and may therefore not be inappropriate. Conversely, if climate data does not have explicit temporal dependencies or trends such as in time-series data, shuffling the data helps randomize the order of the samples during training and can be beneficial for MLPs. This randomization helps to reduce any bias that might arise from the original order of the data and aids in generalizing better to unseen data. This helps the model learn a generalized representation of the data which is not biased by the presentation order of the samples.

2.3 Convolutional neural network (CNN)

CNNs are a type of deep learning neural network designed to automatically learn spatial hierarchies of features from

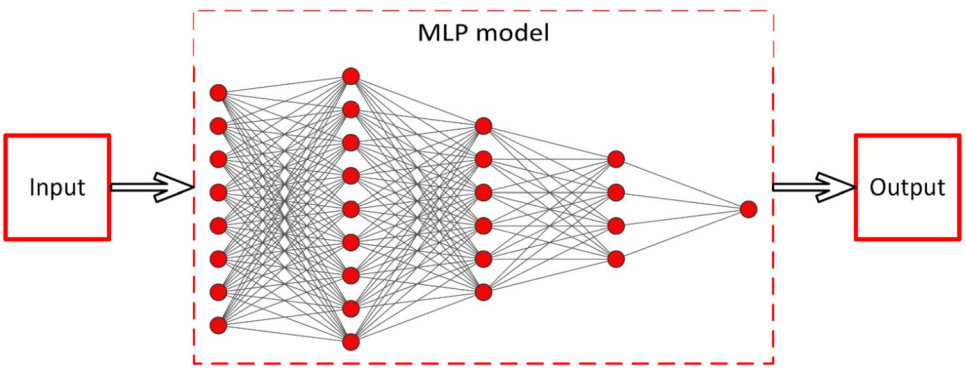

Fig. 4 A schematic structure of an MLP network

input data using convolutional layers. Although CNNs were initially developed for analyzing visual data, they have shown promising results in various fields, including climate science. The ability of CNNs to extract relevant features from large-scale climate data makes them a powerful tool for analyzing complex climate patterns and making accurate predictions [Farhangmehr et al. 2023].

The network structure of the CNN, including convolutional layers, is shown in Fig. 5.

In the current study, a deep three-layer CNN model was used with 64, 32, and 16 kernels. These values achieved the best trade-off between model performance and computational efficiency. Kernel size determines the spatial extent of the local features that the CNN can detect. Hyperparameter tuning led to kernel sizes of 4, 3, and 2 for each convolutional layer.

Maxpooling is a downsampling technique in CNNs that reduces the spatial size of a feature map while preserving crucial information, leading to improved computational efficiency and reduced overfitting risk. The model architecture incorporates a dropout layer with a rate of 0.5, and a max-pooling layer with a size of two, integrated among the convolutional layers.

Optimizers are used to find the optimal set of weights that can be effectively learned from the input data. The RMSprop optimization function with a learning rate of 0.001 was selected for the present model after fine-tuning. The ReLU activation function was chosen through experimentation for its effectiveness in mitigating the vanishing gradient problem during training.

Following the CNN layers, a dense layer with 32 neurons and a rectified linear unit (ReLU) activation function was applied to the model. This architecture allows the model to refine the extracted features and generate the desired output for a given task more effectively. The number of epochs needed for training was 5000 epochs to achieve reasonable performance.

To optimize hyperparameters, a grid search method was employed, exhaustively exploring combinations. Despite potential computational costs, this approach ensures finding the optimal set within the specified grid. For example, in each AI model, four activation functions—ReLU, Sigmoid, tanh, and Linear—were tested, and the best-performing function was selected. Similarly, optimization functions (Adam, SGD, RMSprop) were compared, with the most efficient one chosen. The learning rate was explored within the range of 0.0001 to 0.1. Batch sizes (8, 16, 32, 64) were tested to strike a balance between training time and generalization. The number of epochs was determined using early stopping techniques. The dropout rate varied within its entire range (0 to 1).

2.4 Long short-term memory (LSTM)

Recurrent neural networks (RNN) are a powerful class of neural networks that excel in modeling sequential data. What distinguished them from traditional ANNs is their ability to capture dependencies among consecutive samples. Each data point in a sequence is considered dependent on its predecessors, allowing RNNs to model time-series data effectively. One significant challenge faced by standard RNNs is the vanishing gradient problem, in which gradients diminish as they propagate backward over time. This limitation can hinder the ability of a network to capture long-term dependencies in data. To address this issue, LSTM was introduced as a modified version of RNNs [Tabrizi et al. 2021].

LSTM networks not only overcome the vanishing gradient problem in standard RNNs but also possess the ability to effectively memorize long-term historical information, which makes them highly suitable for modeling and fitting long-term time-series data.

By utilizing specialized memory units, LSTM networks can retain and utilize relevant information from previous time steps, enabling them to capture complex patterns, trends, and dependencies in input data and making them useful for tasks such as time-series forecasting, sequence classification, and sequence generation [Ju and Liu 2021].

An LSTM unit is mainly composed of three gates: an input gate, a forget gate, and an output gate that regulate the flow of information into and out of the cell state [Dai et al. 2023]. The forget gate is used to regulate the amount of information released from the current cell state. In this

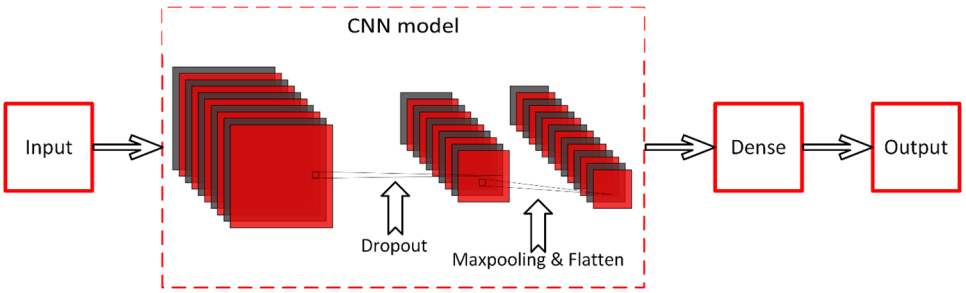

Fig. 5 A schematic structure of a CNN network

gate, an activation function ranges from 0, indicating forgetting everything, to 1, indicating forgetting nothing from the previous timestamps [Babbar et al. 2023]. The input gate controls the flow of current information into the cell state, whereas the output gate determines the amount of information that is communicated to the current output layer.

The network structure of the LSTM including some units is illustrated in Fig. 6.

where x_t and c_t are the input and carry states, respectively, at time step t, and h_{t-1} is the previous hidden state.

The input x_t at time t and the output h_{t-1} of the hidden layer neuron at time $t-1$ are combined and used as the input of the hidden layer at time t, multiplied by different weight vectors [Ju and Liu 2021].

Designing a proper architecture for neural networks plays a crucial role in optimizing the learning process and achieving good performance. In addition, certain hyperparameters must be tuned to achieve an optimal configuration.

In this study, a deep LSTM model with four LSTM layers was employed, each containing 128, 64, 32, and 16 units. Each LSTM unit is crucial for processing and retaining sequential information over time. The number of specified units determines the capacity and complexity of the LSTM layer, which contributes to the performance of the model.

In LSTM networks, the hyperbolic tangent (tanh) activation function is commonly preferred because of its ability to squash output values between -1 and 1. This characteristic enables LSTM to effectively capture and propagate gradients, thereby addressing the vanishing gradient problem. In this study, the tanh activation function was selected for the LSTM model after careful tuning.

A dropout layer with a rate of 0.5 was added after the LSTM layers to address overfitting. Dropout randomly sets a portion of the input units to zero during training, promoting robust learning and reducing the reliance on specific input features. This regularization technique prevents the model from memorizing noise in the training data, leading to a more generalized and accurate model.

Following the LSTM layers, two dense layers were included to process the LSTM output and make the predictions. The first dense layer consisted of 16 neurons with a rectified linear unit (ReLU) activation function, while the final layer contained one neuron with a linear activation function.

The model weights were updated during training using the Adam optimizer with a learning rate of 0.001, after experimenting with various optimizers. The Mean Squared Error (MSE) loss function was employed to quantify the difference between the predicted and true values. Following tuning, a batch size of 32 was selected, representing the number of samples processed in each training iteration. For sufficient learning, the model underwent 500 epochs, signifying the number of iterations of the entire dataset during the training.

The specification of utilized AI models in the present study has been summarized in Table 2.

Each of the employed AI techniques faces limitations in climate modeling, and our efforts have been directed toward enhancing their performance.

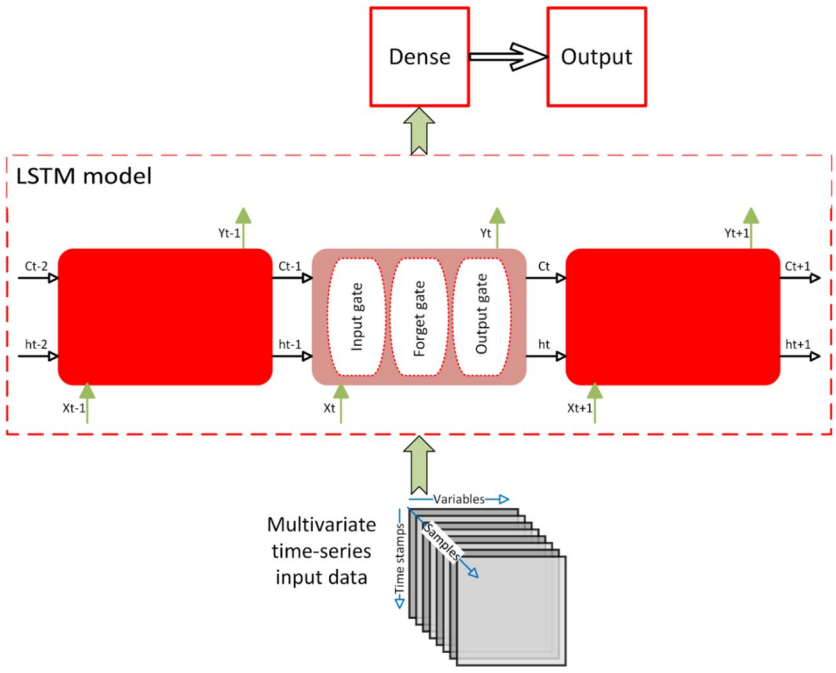

Fig. 6 A schematic structure of an LSTM network

Table 2 The summary of utilized AI models' specifications in the present study

Model	Specifications
MLP	4 dense layers: 128, 64, 32, and 1 neuron Activation function of first three layers: ReLU, last layer: linear Adam optimization function, learning rate: 0.001
CNN	3 convolutional layers: 64, 32, and 16 kernels, kernel sizes: 4, 3, and 2 1 dense layer: 32 neurons, activation function: ReLU A dropout layer, rate of 0.5 A max-pooling layer, size of 2 Optimization function: RMSprop, learning rate: 0.001
LSTM	4 LSTM layers: 128, 64, 32, and 16 units 2 dense layers: 16 neurons, activation function: ReLU; 1 neuron, activation function: linear Activation function: tanh A dropout layer, rate of 0.5 Optimization function: Adam, learning rate: 0.001 Loss function: Mean Squared Error (MSE)

One limitation of the MLP method in climate prediction is its difficulty in capturing complex spatial interactions, a challenge not addressed in this study focused on investigating soil temperature in three different locations. CNNs may exhibit sensitivity to input size and resolution, potentially impacting model performance, especially considering the varied resolutions of climate data. In this study, the inclusion of a Max pooling layer has allowed for the adjustment of input data resolution, improving sensitivity issues related to size and resolution. LSTMs, being prone to overfitting, especially in complex models with limited training data, pose a potential challenge. To address this, regularization techniques were employed in the present study. Additionally, the incorporation of dropout layers was implemented as a strategy to promote a more generalized model.

2.5 Methodological overview

The model input dataset consisted of valid data collected over two consecutive years, 2020 and 2021, comprising approximately 140,000 pieces of climatic information for each location. The output of each model was the predicted hourly soil temperature (K). It is important to note that within each station, the data is evenly distributed over time. This means that we have access to hourly information for all days throughout each year for every climate variable. Table 3 presents a breakdown of the data distribution across different locations and time periods.

For all the models considered, the input data were arranged chronologically and divided into two parts: the training phase comprising the first 65% of the time series, and the remaining 35% used as the testing set. To enhance the model's training stability and address gradient-related challenges, the input data were preprocessed through normalization by subtracting the mean of each feature and dividing it by the standard deviation, resulting in a mean of 0 and a standard deviation of 1 for each feature. Consequently, each model was fed with an 8-dimensional normalized input dataset, while the output consisted of a 1-dimensional soil temperature prediction.

Model performance assessment involves comparing the outputs of the applied AI models. Five error indicators, namely the mean absolute error (MAE), root mean square error (RMSE), coefficient of determination (R-squared), normalized root mean square error (NRMSE), and scatter index (SI), were utilized to measure the quality of the model. The evaluation metrics were defined as [Imanian et al. 2023]

$$MAE = \frac{\sum |y_{obs} - y_{calc}|}{n} \quad \text{Optimal value} = 0 \tag{5}$$

Table 3 Data distribution across different locations and time periods

Station Year	Ottawa	Churchill	Iqaluit
2020 (leap year)	8 × 366 × 24 ≈ 70,000 (variables × days × hours)	8 × 366 × 24 ≈ 70,000	8 × 366 × 24 ≈ 70,000
2021	8 × 365 × 24 ≈ 70,000	8 × 365 × 24 ≈ 70,000	8 × 365 × 24 ≈ 70,000

$$RMSE = \sqrt{\frac{\sum (y_{obs} - y_{calc})^2}{n}} \quad \text{Optimal value} = 0 \tag{6}$$

$$R^2 = \left[\frac{\sum (y_{obs} - \overline{y_{obs}})(y_{calc} - \overline{y_{calc}})}{\sqrt{\sum (y_{obs} - \overline{y_{obs}})^2 \sum (y_{calc} - \overline{y_{calc}})^2}} \right]^2 \quad \text{Optimal value} = 1 \tag{7}$$

$$NRMSE = \frac{RMSE}{[Max(y_{obs}) - Min(y_{obs})]} \quad \text{Optimalvalue} = 0 \tag{8}$$

$$SI = \frac{RMSE}{\overline{y_{obs}}} \quad \text{Optimal value} = 0 \tag{9}$$

where y_{obs} is the observed value; y_{calc} is the interpolated value of the model; $\overline{y_{obs}}$ is the mean of the observed values; $\overline{y_{calc}}$ is the mean of the calculated values; and n is the amount of data.

The deep learning models were implemented using Python 3 programming language with the Keras library, which is a powerful library for developing and evaluating deep learning networks. The cloud-based platform Google Colab, which provides a Google Compute Engine backend with GPU support and RAM of 32 GB, was leveraged.

The overall flow of the simulation is illustrated in Fig. 7.

3 Results

Cross-validation with k-fold divides the dataset into k subsets, training the model k times with different folds as validation sets. This technique provides a more robust estimate of the model's performance, helping overcome overfitting. Overfitting, where a model learns noise in training data that does not generalize well, is mitigated by the comprehensive evaluation of cross-validation. By identifying overfitting early, practitioners can choose models that generalize effectively to new data, improving reliability and predictive capabilities. The dataset is split into training and validation sets for each fold, and RMSE is calculated as a performance metric across all folds. The average and standard deviation of the RMSE from 5-folds are reported in Table 4.

As discussed in Sect. 2, the study incorporated hourly data over the entire year from eight climate variables: air temperature, precipitation, surface pressure, evaporation, wind gusts, dew point temperature, solar radiation, and thermal radiation. The objective was to predict the hourly soil temperature, with distinct predictions for 2020 and 2021. The investigation was conducted across three stations representing diverse cold climates: Ottawa for cold, Churchill for very cold, and Iqaluit for polar climates. Employing a machine learning framework, the prediction was facilitated by Convolutional Neural Networks (CNN), long short-term memory (LSTM), and multilayer perceptron (MLP) methods, as discussed in Sect. 2. The input data used for these methods were chronologically arranged. Furthermore, the MLP model was applied with

Table 4 Error analysis of the employed models

Models	CNN	LSTM	MLP	MLP-shuffled
Mean fivefold RMSE	3.29	2.73	2.72	1.74
St dev fivefold RMSE	1.15	0.97	1.20	0.12

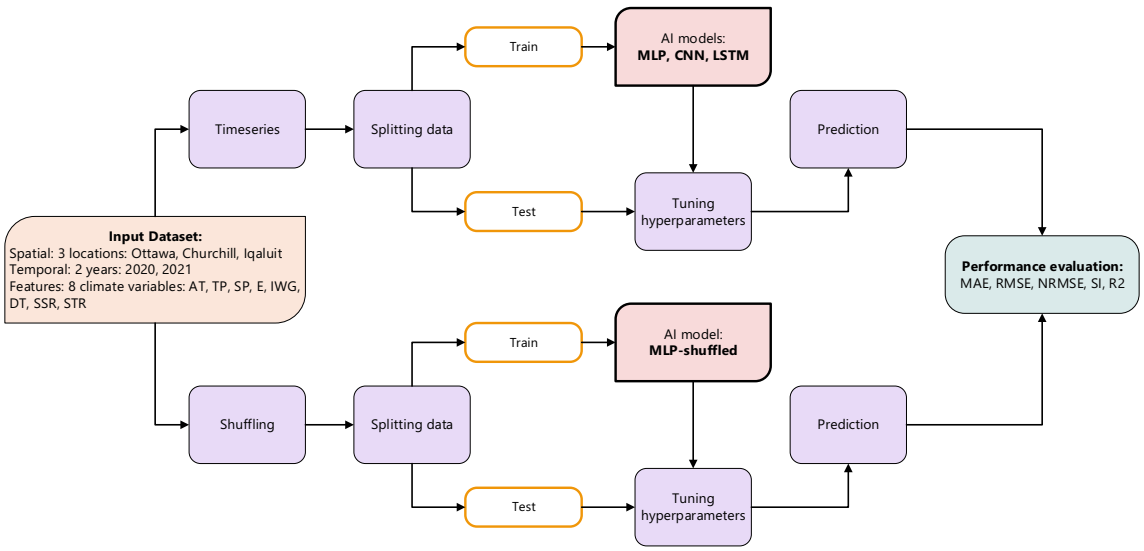

Fig. 7 Algorithm of the developed numerical model in the present study

shuffled input data to assess the influence of data order on model performance. Hence, four machine learning approaches were employed to forecast soil temperature across three stations during two distinct time periods, 2020 and 2021 respectively.

The results of RMSE and *R*-squared using Eqs. 6 and 7, in the training and testing stage for each model are presented in Table 5.

Assessing the model's performance with appropriate error metrics provides valuable insights, enhancing its robustness. Results of the error analysis using Eqs. 5 to 9 are summarized in Table 6.

4 Discussion

The outcomes of the error analysis for soil temperature prediction, computed via various machine learning techniques and employing the error metrics are depicted in Fig. 8. The comparison of error indicators provides insights into the performance of each model in accurately forecasting soil temperature at different locations.

The first two error indicators depicted in Fig. 8a and b are MAE and RMSE, both of which have the same unit as the calculated data, namely, Kelvin. MAE provides the average absolute magnitude of the errors without considering their direction. In Fig. 8, it is evident that the CNN model exhibits a higher MAE, indicating lower accuracy, while the MLP-shuffled model performs remarkably well, showing the best performance in soil temperature prediction among all cases with a lower MAE. In contrast, RMSE provides another perspective on the average magnitude of the errors. In MAE, all individual differences are treated equally, whereas RMSE assigns a greater weight to larger errors. Consistently, the MLP-shuffled method had the lowest RMSE values, indicating a higher prediction accuracy than the other methods. The LSTM and MLP methods exhibited fairly comparable performance, with LSTM outperforming MLP in certain cases, whereas MLP showed a slight advantage in others.

The difference between the RMSE and MAE indicates the variability in the errors. With differences of 0.75, 0.65, 0.55, and 0.45 Kelvin for the CNN, LSTM, MLP, and MLP-shuffled, respectively, the MLP-shuffled models emerge as the one with the least error variance in soil temperature prediction.

NRMSE and SI are indices derived from RMSE but are dimensionless, which provides a better understanding for comparison. The NRMSE normalizes the RMSE by the range of observations to provide a relative measure of error that is independent of the scale of the data. By dividing the RMSE by the range of observed values, it essentially scales the RMSE to the same unit as the data, making it easier to compare across different datasets with varying ranges. The scatter index, on the other hand, normalizes the RMSE by the average of the observed values. This index was used to assess the relative dispersion of predicted values around the mean of the observed data. Similar to the normalized RMSE, this metric also allows for comparisons across datasets with different scales.

In contrast to MAE and RMSE, which serve to compare the performance of various ML methods, NRMSE and SI

Table 5 Performance of models in training and testing stages

Model	Error metrics	Station/year	Ottawa, 2020	Ottawa, 2021	Churchill, 2020	Churchill, 2021	Iqaluit, 2020	Iqaluit, 2021
CNN	RMSE (K)	Training	1.58	1.54	0.89	1.34	1.43	0.95
		Testing	2.52	2.34	4.02	2.77	3.29	1.81
	R-squared	Training	98.69%	98.10%	98.49%	98.38%	98.49%	98.22%
		Testing	90.25%	91.42%	85.82%	91.50%	61.24%	76.73%
LSTM	RMSE (K)	Training	1.61	3.00	2.18	3.30	1.33	1.16
		Testing	2.11	2.35	2.47	4.16	1.99	1.92
	R-squared	Training	97.60%	97.16%	97.09%	91.58%	97.82%	97.48%
		Testing	91.79%	93.28%	93.15%	87.89%	73.95%	78.31%
MLP	RMSE (K)	Training	1.88	1.67	1.88	2.52	1.40	1.07
		Testing	2.16	2.72	2.37	3.26	1.54	1.55
	R-squared	Training	96.87%	97.09%	97.80%	94.82%	97.56%	97.85%
		Testing	92.62%	91.05%	93.75%	92.59%	67.64%	86.13%
MLP-shuffled	RMSE (K)	Training	1.79	2.00	2.01	2.37	1.67	1.42
		Testing	1.78	2.03	2.01	2.41	1.68	1.46
	R-squared	Training	96.74%	95.69%	97.03%	95.62%	95.55%	95.06%
		Testing	96.71%	95.56%	97.09%	95.61%	95.59%	94.98%

Table 6 Error analysis of the employed models

Station/year	Model Error metrics	CNN	LSTM	MLP	MLP-shuffled
Ottawa, 2020	MAE (K)	1.93	1.643211	1.705694	1.44942
	RMSE (K)	2.52	2.287617	2.262778	1.964682
	NRMSE	0.06894	0.062583	0.061903	0.053748
	SI	0.009028	0.008195	0.008106	0.007038
	R-squared	0.9025	0.917123	0.918097	0.960273
Ottawa, 2021	MAE (K)	1.75	1.520972	1.718574	1.433457
	RMSE (K)	2.34	2.090795	2.354091	1.965223
	NRMSE	0.056635	0.050603	0.056976	0.047564
	SI	0.008357	0.007467	0.008407	0.00707
	R-squared	0.9142	0.946154	0.932826	0.958194
Churchill, 2020	MAE (K)	2.403223	1.876665	2.437716	1.542529
	RMSE (K)	3.741217	2.470424	2.95455	1.975654
	NRMSE	0.094944	0.062694	0.07498	0.050138
	SI	0.013889	0.009171	0.010968	0.007334
	R-squared	0.831898	0.931513	0.901163	0.971556
Churchill, 2021	MAE (K)	2.396128	3.014004	2.511985	1.789534
	RMSE (K)	3.010114	4.159305	3.564946	2.377223
	NRMSE	0.063259	0.08741	0.074919	0.049959
	SI	0.011109	0.01535	0.013157	0.008773
	R-squared	0.88317	0.878892	0.911112	0.956462
Iqaluit, 2020	MAE (K)	2.566779	1.551522	1.208466	0.966121
	RMSE (K)	3.285559	1.993165	1.537967	1.461451
	NRMSE	0.141106	0.085601	0.066051	0.062765
	SI	0.012088	0.007333	0.005658	0.005377
	R-squared	0.612365	0.739492	0.676448	0.966121
Iqaluit, 2021	MAE (K)	1.317409	1.488145	1.208466	0.966064
	RMSE (K)	1.814137	1.918807	1.537967	1.182501
	NRMSE	0.091623	0.096909	0.077675	0.059722
	SI	0.006706	0.007093	0.005685	0.004371
	R-squared	0.76734	0.783117	0.852897	0.966064

were employed to assess the performance of ML models in predicting soil temperature across different stations, representing distinct climates. Because the range and mean of soil temperature at these stations vary, analyzing the NRMSE and SI indices provides a clearer understanding of which models excel in specific cold climates.

Figure 8c and d reveals that the error metrics for station Ottawa were notably lower than those for the other two stations. Across all prediction methods, higher errors were observed for colder locations, namely, Churchill and Iqaluit. For instance, the NRMSE index in Fig. 8d indicates that the weakest performance was observed in the case of the Iqaluit station with a polar climate.

Nevertheless, MLP-shuffled consistently outperformed other ML methods employed for predicting soil temperature. In summary, it can be concluded that the ML methods employed struggle to deliver precise predictions of soil temperature in very cold and polar climates. On the other hand, alleviating the time dependency in the data by using shuffled datasets consistently improves the capability of the network to forecast soil temperature, as demonstrated by all metrics.

In Fig. 8, the MLP-shuffled model shows exceptional performance, surpassing other models that consider time-series data. It attains the lowest error indicators and highest *R*-squared values, underscoring its robust predictive capabilities for soil temperature. Comparatively, the LSTM model outperformed the CNN and MLP models with time-series data, by effectively capturing the intrinsic time dependencies and historical patterns in the time-series data, resulting in more accurate predictions. Its ability to retain long-term dependencies makes it particularly well-suited for tasks

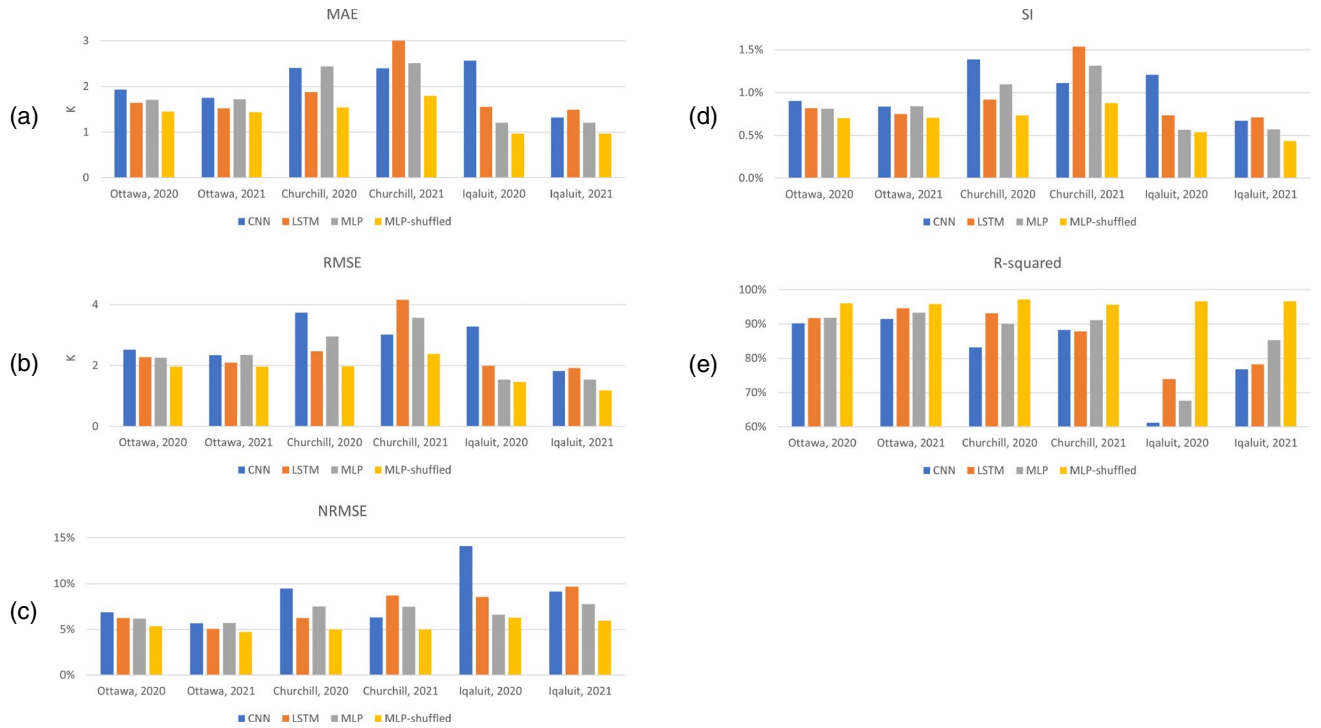

Fig. 8 Evaluation indicators calculated in different stations using various methods: **a** mean absolute error, **b** root mean square error, **c** normalized root mean square error, **d** scatter index, and **e** *R*-squared

such as climate prediction, where sequential data are pivotal. However, in the context of soil temperature prediction, where current climate conditions, primarily air temperature, have a greater impact, as evidenced by the correlation matrix in Fig. 3, the MLP model, which handles shuffled variables, delivered superior error indicators.

To further support this finding, Fig. 9 presents the time series data of soil temperature and air temperature for the Iqaluit station in 2020 and 2021. It is worth noting that the soil temperature (ST) and air temperature (AT) exhibited a strong correlation over the warmer part of the year, as shown in Fig. 3. Figure 9 shows that prior to the onset of the frost season, variations in air and soil temperatures were closely aligned. However, during the cold season, when the air temperature drops well below freezing, the soil temperature remains near the freezing point and differs from the earlier pattern. Furthermore, although air temperature can undergo rapid fluctuations, soil temperature changes are delayed, and they rarely descend below the freezing point when the ground is covered by an ice layer. This phenomenon helps explain why ML models struggle more to accurately predict the soil temperature in polar regions, as depicted in Fig. 8 with larger discrepancies in the metrics for the Iqaluit area. The temperature data for the first three cold months (January, February, and March) and the last two cold months (November and December) are highlighted on the graphs in Fig. 9.

The scatter plots in Fig. 10 illustrate the soil temperature values predicted by the MLP-shuffled model compared to the actual data. The close alignment of the regression line with the identity line highlights a strong fit between the observed values and the model predictions. These findings emphasize the reliability of the MLP-shuffled model for delivering accurate soil temperature predictions across different stations and climates.

The manuscript significantly contributes to the application of ML in forecasting soil parameters, focusing on cold climates. The study's contributions to knowledge encompass:

The challenges in cold climates, where ML faces difficulties accurately predicting soil temperature due to the unique characteristics of ice-covered soil. Historical trends alone prove insufficient for reliable forecasting.

The superiority of the introduced MLP-shuffled model, which operates with shuffled input data and outperforms

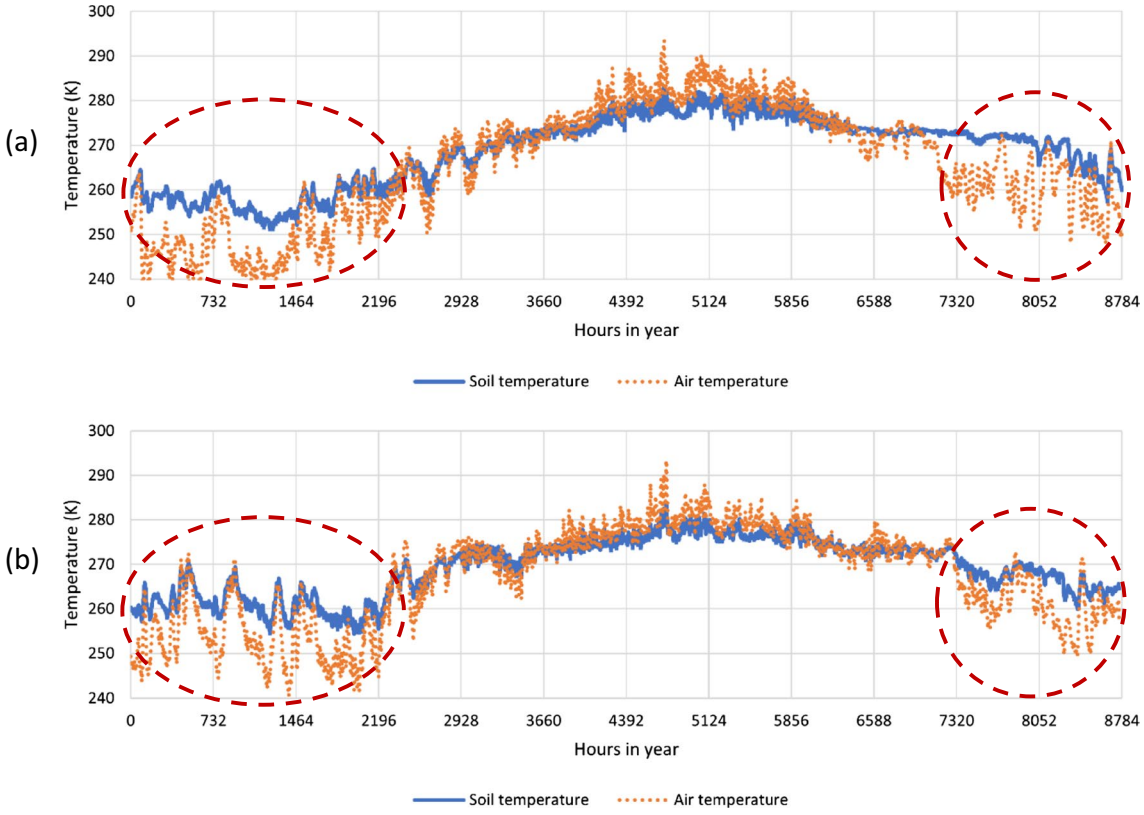

Fig. 9 Time series of air temperature and soil temperature in Iqaluit station: **a** 2020 and **b** 2021. The data for the Jan–Mar and Nov–Dec periods are specified with red dashed lines

other methods. This demonstrates superior prediction accuracy, attributed to the model's capacity to acquire a generalized data representation without biases from the sample presentation sequence.

Practical implications for climate forecasting, as the study's insights underscore the potential of ML techniques, especially MLP-shuffled, in enhancing our understanding and predictive capabilities in climate-related research.

5 Conclusion

A precise and cost-effective model empowered by machine learning techniques was developed for soil temperature forecasting in cold climates, studying six cases of Ottawa-2020, Ottawa-2021, Churchill-2020, Churchill-2021, Iqaluit-2020, and Iqaluit-2021. To comprehensively evaluate the predictive approaches for soil temperature, four ML models—CNN, LSTM, MLP in time series, and MLP on shuffled data—were employed. Utilizing eight hourly land and atmospheric variables, including air temperature, precipitation, surface pressure, evaporation, wind gust, dew point temperature, solar radiation, and thermal radiation as the input dataset, predictions were made across various stations in Canada, each representing a different cold climate.

Our findings indicate that ML is promising as an effective approach for forecasting soil parameters. The models considered demonstrated a reliable capacity for soil temperature prediction in cold climates, emphasizing the overall potential of machine learning techniques in enhancing our understanding and predictive capabilities in the scope of climate forecasting.

The main findings of this study are outlined as follows:

– Machine learning methods, when using input data as time series, struggle to predict soil temperature accurately, especially in very cold and polar climates. This is because soil temperature cannot be reliably forecast based solely on the historical trends of land and atmospheric variables.

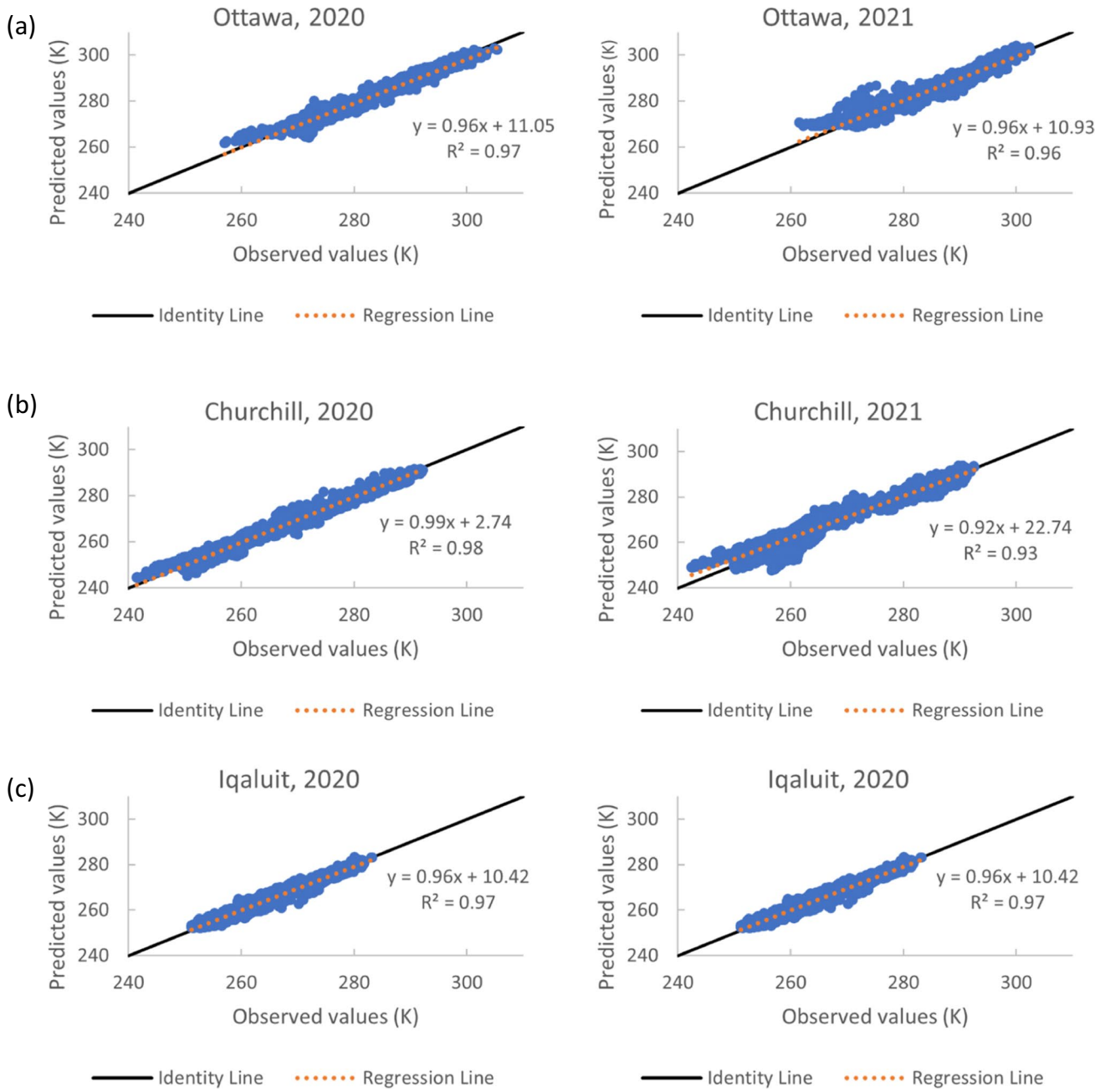

Fig. 10 Scatter plots of predicted and observed soil temperature in different stations using shuffled MLP method: **a** station Ottawa, **b** station Churchill, **c** station Iqaluit

In cold climates with ice-covered soil, the soil temperature remains close to the frost point even when the air temperature drops below freezing. This underscores the need for models that can accommodate these distinctive characteristics under extremely cold conditions.

– MLP-shuffled, which operates with shuffled input data, distinguishes itself by yielding the lowest error metrics, signifying superior prediction accuracy compared to other methods. This is primarily attributed to the capacity of the model to acquire a generalized data representation, devoid

of bias introduced by the sequence in which samples are presented.

This study primarily utilized common climate variables applicable to all regions. Future research can explore the impact of specialized cold region climate variables, such as snow depth and snow density, on soil temperature prediction in cold climates.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by HI, AM, and VF. The first draft of the manuscript was written by HI. All authors commented on previous versions of the manuscript and all authors read and approved the final manuscript.

Funding This research was funded by the National Research Council Canada through the Artificial Intelligence for Logistics Supercluster Support Program, grant number AI4L-120.

Data availability The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

References

Babbar N, Kumar A, Kuma Verma V (2023) Forecasting wheat yield using long short-term memory considering soil and metrological parameters. 3rd Int Conf Int Commun Comput Tech ICCT 2023 Jaipur, India. IEEE pp 1-4. https://doi.org/10.1109/ICCT56969.2023.10076090

Bayatvarkeshi M, Bhagat SK, Mohammadi K, Kisi O, Farahani M, Hasani A, Deo R, Yaseen ZM (2021) Modeling soil temperature using air temperature features in diverse climatic conditions with complementary machine learning models. Comput Electron Agric 185:106158. https://doi.org/10.1016/j.compag.2021.106158

Beck HE, Zimmermann NE, McVicar TR, Vergopolan N, Berg A, Wood EF (2018) Present and future Köppen-Geiger climate classification maps at 1-km resolution. Sci Data 5:180214. https://doi.org/10.1038/sdata.2018.214

Belouz K, Zereg S (2023) Extreme learning machine for soil temperature prediction using only air temperature as input. Environ Monit Assess 195(8):962. https://doi.org/10.1007/s10661-023-11566-2

Bonakdari H, Moeeni H, Ebtehaj I, Zeynoddin M, Mahoammadian A, Gharabaghi B (2019) New insights into soil temperature time series modeling: Linear or nonlinear? Arch Meteorol Geophys Bioclimatol Ser B 135:1157–1177

Dai B, Yang W, Ji X, Zhu F, Fang R, Zhou L (2023) An ensemble deep learning model for short-term road surface temperature prediction. J Trans Eng Part B: Pavements 149:1. https://doi.org/10.1061/jpeodx.pveng-1192

Delbari M, Sharifazari S, Mohammadi E (2019) Modeling daily soil temperature over diverse climate conditions in Iran—a comparison of multiple linear regression and support vector regression techniques. Theoret Appl Climatol 135(3–4):991–1001. https://doi.org/10.1007/s00704-018-2370-3

Ebtehaj I, Bonakdari H, Samui P, Gharabaghi B (2023) Multi-depth daily soil temperature modeling: meteorological variables or time series? Theoret Appl Climatol 151(3–4):989–1012. https://doi.org/10.1007/s00704-022-04314-y

Farhangmehr V, Cobo JH, Mohammadian A, Payeur P, Shirkhani H, Imanian H (2023) A convolutional neural network model for soil temperature prediction under ordinary and hot weather conditions: comparison with a multilayer perceptron model. Sustainability 15:7897. https://doi.org/10.3390/su15107897

Hao H, Yu F, Li Q (2021) Soil Temperature Prediction Using Convolutional Neural Network Based on Ensemble Empirical Mode Decomposition. IEEE Access 9:4084–4096

Hou J, Wang Y, Zhou J, Tian Q (2022) Prediction of hourly air temperature based on CNN–LSTM. Geomat Nat Haz Risk 13(1):1962–1986. https://doi.org/10.1080/19475705.2022.2102942

Imanian H, Hiedra Cobo J, Payeur P, Shirkhani H, Mohammadian A (2022) A comprehensive study of artificial intelligence applications for soil temperature prediction in ordinary climate conditions and extremely hot events. Sustainability 14:8065. https://doi.org/10.3390/su14138065

Imanian H, Shirkhani H, Mohammadian A, Hiedra Cobo J, Payeur P (2023) Spatial interpolation of soil temperature and water content in the land-water interface using artificial intelligence. Water 15:473. https://doi.org/10.3390/w15030473

Ju J, Liu FA (2021) Multivariate time series data prediction based on ATT-LSTM network. Appl Sci (Switzerland) 11:20. https://doi.org/10.3390/app11209373

Mehdizadeh S, Behmanesh J, Khalili K (2018) Comprehensive modeling of monthly mean soil temperature using multivariate adaptive regression splines and support vector machine. Arch Meteorol Geophys Bioclimatol Ser B 133:911–924

Parveen N, Zaidi S, Danish M (2020a) Artificial intelligence (AI)-based friction factor models for large piping networks. Chem Eng Commun 207(2):213–230. https://doi.org/10.1080/00986445.2019.1578757

Parveen N, Zaidi S, Danish M (2020b) Comparative analysis for the prediction of boiling heat transfer coefficient of R134a in micro/mini channels using artificial intelligence (AI)-based techniques. Int J Model Simul 40(2):114–129. https://doi.org/10.1080/02286203.2018.1564809

Pomeroy JW, Brown T, Fang X, Shook KR, Pradhananga D, Armstrong R, Harder P, Marsh C, Costa D, Krogh SA, Aubry-Wake C, Annand H, Lawford P, He Z, Kompanizare M, Lopez Moreno JI (2022) The cold regions hydrological modelling platform for

hydrological diagnosis and prediction based on process understanding. J Hydrol 615(A):128711. https://doi.org/10.1016/j.jhydrol.2022.128711

Samadianfard S, Ghorbani MA, Mohammadi B (2018) Forecasting soil temperature at multiple-depth with a hybrid artificial neural network model coupled-hybrid firefly optimizer algorithm. Inf Process Agric 5:465–476

Seifi A, Ehteram M, Nayebloei F, Soroush F, Gharabaghi B, Haghighi AT (2021) GLUE uncertainty analysis of hybrid models for predicting hourly soil temperature and application wavelet coherence analysis for correlation with meteorological variables. Soft Comput 25:10723–10748

Tabrizi SE, Xiao K, Van Griensven Thé J, Saad M, Farghaly H, Yang SX, Gharabaghi B (2021) Hourly road pavement surface temperature forecasting using deep learning models. J Hydrol 603:126877. https://doi.org/10.1016/j.jhydrol.2021.126877

Taheri M, Schreiner HK, Mohammadian A, Shirkhani H, Payeur P, Imanian H, Cobo JH (2023) A review of machine learning approaches to soil temperature estimation. Sustainability 15:7677. https://doi.org/10.3390/su15097677

Xing L, Li L, Gong J, Ren C, Liu J, Chen H (2018) Daily soil temperatures predictions for various climates in United States using data-driven model. Energy 160:430–440. https://doi.org/10.1016/j.energy.2018.07.004

Yener D, Ozgener O, Ozgener L (2017) Prediction of soil temperatures for shallow geothermal applications in Turkey. Renew Sustain Energy Rev 70:71–77. https://doi.org/10.1016/j.rser.2016.11.065

Yu L, Liu Y, Bu K, Wang WJ, Zhang S (2022) Soil temperature mitigation due to vegetation biophysical feedbacks. Global Planet Change 218:103971. https://doi.org/10.1016/j.gloplacha.2022.103971

Zeynoddin M, Bonakdari H, Gumiere SJ, Rousseau AN (2023) Multi-tempo forecasting of soil temperature data; application over Quebec. Can Sustain 15(12):9567. https://doi.org/10.3390/su15129567