Received: 13 June 2025 | Revised: 27 August 2025 | Accepted: 21 November 2025 | Published online: 14 January 2026

RESEARCH ARTICLE

Journal of Data Science and Intelligent Systems
2026, Vol. 00(00) 1–5
DOI: 10.47852/bonviewJDSIS62026463

BON VIEW PUBLISHING

AI-Powered Soil Temperature Modeling for Sustainable Agriculture in Arid Regions: A Case Study of Bustan, Uzbekistan

Lakindu Mampitiya[1,2], Namal Rathnayake[3], Kenjabek Rozumbetov[4,5], Valery Erkudov[6], Mirzohid Koriyev[7], Komali Kantamaneni[8] and Upaka Rathnayake[9,*]

[1] *Department of Mechanical Engineering, University of Sri Jayewardenepura, Sri Lanka*

[2] *Water Resources Management and Soft Computing Research Laboratory, Sri Lanka*

[3] *Advanced Institute for Marine Ecosystem Change, Japan Agency for Marine-Earth Science and Technology, Japan*

[4] *Department of Veterinary Diagnostics and Food Safety, Nukus Branch Of The Samarkand State University Of Veterinary Medicine, Uzbekistan*

[5] *Department of General Biology and Physiology, Karakalpak State University, Uzbekistan*

[6] *Department of Normal Physiology, Saint-Petersburg State Pediatric Medical University, Russia*

[7] *Department of Natural Sciences, Namangan State Pedagogical Institute, Uzbekistan*

[8] *School of Engineering, University of Central Lancashire, UK*

[9] *Department of Civil Engineering & Construction, Atlantic Technological University, Ireland*

Abstract: Soil temperature is a key determinant of soil health and agricultural productivity, especially in arid regions vulnerable to climate change. This study investigates the use of advanced machine learning models to predict soil temperature variations in Bustan, Uzbekistan, a region facing significant climatic stress. Using 16 years of meteorological data, including atmospheric temperature, humidity, and wind speed, eight machine learning models were evaluated for their ability to predict surface and subsurface (10 cm depth) soil temperatures. Among the models tested, the bi-directional long short-term memory (Bi-LSTM) algorithm demonstrated superior predictive accuracy with R^2 values exceeding 0.94 for subsurface temperatures. The two-step modeling approach utilized Bi-LSTM outputs from surface temperature predictions to inform subsurface estimates, reflecting a novel methodology for climate-sensitive agriculture. The results provide a practical framework for improving irrigation planning, crop yield forecasting, and sustainable land management in data-scarce arid environments. By demonstrating high accuracy and real-world applicability, this AI-driven model offers a scalable solution for enhancing agricultural resilience in Uzbekistan and similar contexts.

Keywords: arid climate modeling, Bi-LSTM, machine learning, soil temperature prediction, sustainable agriculture

1. Introduction

Soil plays an important role in maintaining ecological balance and sustaining life [1, 2]. It consists of matters in all three states (solid, liquid, and gas), and has numerous contributions to the natural systems that make it essential for the functioning of the Earth. The composition of soil includes all three aggregate states and is chemically enriched with environmentally friendly elements such as oxygen, silicon, aluminum, nitrogen, phosphorus, potassium, calcium, magnesium, carbon, and hydrogen. Soil, therefore, plays a decisive role in stabilizing the atmosphere, the lithosphere, and the hydrosphere. It supports critical processes such as nutrient cycling, microbial activity, and ecological balance [3, 4]. Soil cover is crucial for maintaining the balance of natural cycles on a global scale (soil and water) [5, 6].

For plants, the soil serves as a nutrient and water reservoir, supports root systems, and facilitates gas exchange, a crucial function for plant respiration [7, 8]. For animals, it provides habitats for nesting and burrowing and food sources and also contributes to the overall functioning of terrestrial ecosystems [9]. Microbes, including bacteria, fungi, and actinomycetes, play an essential role in the decomposition of organic matter and enrichment of the soil [10, 11], thereby essential for the existence of both plants and animals. Additionally, soil contributes significantly to human existence by supporting agriculture and ecological balance [12, 13].

In addition, soil quality is a crucial factor that directly and indirectly influences climatic conditions [14, 15]. Conversely, changing climate patterns, such as irregular rainfall and temperature fluctuations, have a significant impact on soil conditions [16, 17]. These changes are often caused due to human activities that disrupt the ecological balance.

Soil temperature, a key factor in soil functionality, is influenced by climatic conditions such as varying precipitation, air pollution, and

***Corresponding author:** Upaka Rathnayake, Department of Civil Engineering & Construction, Atlantic Technological University, Ireland. Email: upaka.rathnayake@atu.ie

temperature fluctuations [18–20]. Changes in soil temperature can influence microbial activity [21], nutrient cycling [22], and plant growth [23]. Therefore, the changes in soil temperatures can significantly affect livelihood. The disruption in the ecological system should be taken seriously when working toward a better future.

Agricultural food production is one of the global issues with grave impact due to climate change [24–27]. As demand continues to grow, the damage caused by various adverse climate events presents a significant challenge to agricultural food production. Consequently, many countries are struggling to address this issue effectively [28], mainly due to instant but extreme climatic events. However, a slow but steady impact on agriculture is observed owing to notable change in soil temperatures [29–31], which is not only the surface temperatures of the soil layer but also in the inner soil layers.

The topsoil layer (0–10 cm), which interacts most with the external environment, is particularly susceptible to the fluctuations of soil temperatures [32, 33]. To mitigate the impact of changing soil temperatures, it is crucial to understand and control these fluctuations. Proper analysis and treatment of the affected soil layers can help restore natural ecosystems. However, this depends heavily on the quality of data collection and analysis [34, 35].

Uzbekistan, a Central Asian country, faces significant challenges related to soil temperature fluctuations. Its geographical location, characterized by remoteness and distance from large bodies of water, affects the accumulation of minerals in the soil. Problems such as soil salinization and water scarcity, which are exacerbated by Uzbekistan's dry climate, further complicate the situation [36].

Fluctuations in the soil temperature also have a negative impact on sustainability, quality of life, and agricultural productivity. Overcoming these challenges is crucial for the development of Uzbekistan as a country dependent on agriculture, as these soil problems could lead to economic crises. Therefore, predicting soil temperatures both on surface and inner soil layers are highly important.

Several attempts have been made to predict soil temperatures using meteorological data with the aid of machine learning in many areas of the world [37–43]. However, only Mampitiya et al. [32] evaluated the soil temperatures, but only in Nukus, Uzbekistan.

While extensive research has highlighted the critical roles of soil in ecological balance, plant and microbial life, and climate regulation, there remains a significant gap in understanding the dynamic interplay between soil temperature fluctuations and their impact on soil functionality, particularly in arid and semi-arid regions like Uzbekistan. Although global studies have examined the effects of climate change on soil processes, they often generalize findings without accounting for regional variability in soil structure, depth-specific temperature sensitivity, and localized climate conditions. Moreover, much of the existing literature emphasizes surface soil layers, overlooking the deeper layers that are equally affected by long-term temperature shifts. The slow yet profound impact of inner soil layer temperature changes on microbial activity, nutrient cycling, and agricultural productivity is underexplored. In regions like Uzbekistan, where agriculture is highly climate-sensitive, and soil degradation is accelerated by salinization and water scarcity, the lack of precise, high-resolution data on soil temperature dynamics limits the development of targeted mitigation strategies. Addressing this knowledge gap is essential for improving soil management practices, ensuring food security, and enhancing climate resilience in vulnerable agro-ecological systems. Therefore, in this study, a predictive analysis of soil temperature in Bustan, Uzbekistan, was conducted using advanced artificial intelligence models. The research aims to provide actionable insights to control and treat soil temperature and help Uzbekistan develop effective action plans to address these challenges. In this way, this work will contribute to sustainable agriculture and improved living standards in Uzbekistan.

2. Study Area and Data

2.1. Case study area

Bustan (41.8455° N, 60.9169° E), a city in Uzbekistan, was selected as the case study area for this application. Bustan is of great importance due to its geographical location, climatic conditions, and agricultural importance, making it a valuable case study. Bustan is the administrative center of Ellikqala district in Karakalpakstan, Uzbekistan. Its location near bodies of water and its climatic and agricultural characteristics make it unique in the region. Figure 1 presents the case study area.

Bustan is located in the southern part of Uzbekistan, close to the national border. The bodies of water such as the Amu Darya River (within a 50 km radius) and Lake Akhchakol (within a 10 km radius) enrich the surrounding region. These water resources play a crucial role in supporting agriculture and maintaining soil moisture in the area [44]. Compared to other parts of Uzbekistan, Bustan is better endowed with water sources, which increases agricultural productivity.

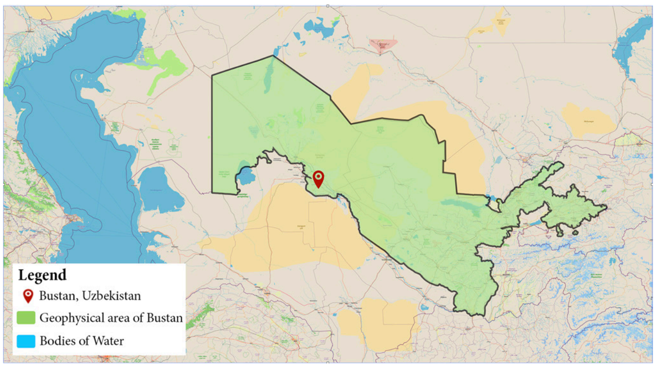
The region is also located in a non-tropical desert zone with unique vegetation. The Ustyurt Plateau, the Amu Darya Delta, the northwestern Kyzylkum Desert, and low mountainous zones characterize its natural geography. In particular, the Ustyurt Plateau is characterized by gypsum, salt, and sand deserts with vegetation consisting of gypsophytes, halophytes, and psammophytes. These diverse landscapes reflect the complexity of the region's ecological and geographical systems [45].

Agriculture is the backbone of Bustan's economy, with cotton being the most important crop. Nearby water sources greatly support the region's agricultural systems, providing better soil maintenance and higher nutrient levels than in other parts of Uzbekistan [46]. These conditions also promote rich microbial activity in the soil, which is crucial for sustaining crop production. The fertile soil and water availability make Bustan an agricultural center of the country.

Unique climatic conditions prevail in Bustan. Daytime temperatures rise significantly, while nighttime temperatures fall relatively little, resulting in significant temperature fluctuations. Rainfall is unevenly distributed throughout the year, with months of little or no rainfall followed by a gradual increase in the middle of the year, which peaks before decreasing again [47]. Humidity in Bustan varies due to its proximity to bodies of water and desert areas. For example, from June to August, the city heavily experiences desert conditions that increase temperature, which are exacerbated by wind conditions. The interplay of these climatic factors creates a complex and dynamic environment in Bustan.

The soil in the Bustan region has different characteristics due to its geography and climate. The Amu Darya delta is dominated by alluvial meadows, forest soils, and dry lake beds. The Ustyurt Plateau has barren Shorhok (salty) soils, while grey-brown soils predominate

Figure 1
Geographical location of Bustan, Uzbekistan



in Kyziljar and Tokmoktog, and sandy desert soils can be found in the dry soil of the Aral Sea [32, 48]. Soil temperatures in Bustan fluctuate throughout the year and are influenced by seasonal changes and climatic conditions. Above and below-ground layers experience alternating rises and falls in temperature. These fluctuations are closely related to seasonal shifts and the influence of nearby bodies of water, which directly affect soil moisture and nutrient levels. Seasonal temperature fluctuations increase microbial activity and nutrient cycling in warmer months, but slow them down during colder months. These temperature fluctuations are mitigated by near bodies of water as they maintain higher soil moisture, which is essential for agriculture.

2.2. Data collection

Soil temperatures at the surface and 10 cm below the surface, atmospheric temperatures (minimum, maximum, and mean), relative humidity (mean value, minimum value), and wind speed (m/s) were collected from the Meteorological Center of the Republic of Karakalpakstan, Uzbekistan. Data were collected covering 16 years from 2008 to 2023. The meteorological center has recorded three data samples for each month. Therefore, there are 576 data points in the analysis. Of the parameters influential to the soil temperature, precipitation was not included in this analysis due to its non-significant occurrence. The area was not receiving significant rainfall (on average 2–8 mm/month) (rainfall statistics at https://www.weather-atlas.com/en/uzbekistan/bustan-climate).

3. Machine Learning Model Algorithms Used

In this study, eight machine learning algorithms were used for predictions. A brief description of each machine learning algorithms used are given below.

1) XGBoost is a highly efficient algorithm known for its speed, accuracy, and error minimization using the decision tree-based ensemble methods. It outperforms traditional algorithms such as Random Forest, support vector machines (SVMs), and neural networks and shows exceptional predictive power in various domains. For example, XGBoost has achieved 98.49% accuracy in bioactivity prediction [49], 99.9% accuracy in material science applications [50], and superior performance in urban land use classification [51]. In soil studies, XGBoost has shown high accuracy in predicting soil salinity [52] and soil organic matter using remote sensing data [53]. Its ensemble learning approach ensures robust predictions while maintaining interpretability, making it invaluable for geotechnical and environmental modeling [54].

2) CatBoost specializes in the efficient handling of categorical data and offers advanced feature selection mechanisms and reduced computational costs. It is ideally suited for environmental applications such as soil moisture and precipitation prediction. Studies highlight CatBoost's ability to integrate atmospheric and soil parameters, achieving R^2 values of up to 0.9935 for soil moisture prediction [55]. It also outperforms algorithms such as XGBoost and ridge regression in precipitation modeling and achieves high accuracy metrics for daily and weekly forecasts [56]. This adaptability to complex temporal and spatial patterns makes CatBoost a reliable tool for agricultural water management and urban meteorology [57].

3) Long short-term memory (LSTM) networks, a class of recurrent neural networks (RNNs), are characterized by the recognition of sequential data patterns and long-term dependencies. They have been used extensively in hydrology and have outperformed models such as SAC-SMA and SWAT in predicting precipitation and runoff [58, 59]. Their interpretability is consistent with hydrological principles and reveals insights into the dynamics of water storage, such as soil moisture and snow [60]. Despite challenges such as data scarcity and limitations in predicting low flows, the integration of physical principles such as mass balance increases their reliability in unconfined catchments [61].

4) Bidirectional LSTM (Bi-LSTM) extends its capabilities by processing data bidirectionally and capturing both past and future dependencies. This bidirectional architecture has demonstrated superior accuracy in hydrology and soil science, particularly in the modelling of precipitation and runoff. For example, Bi-LSTM models augmented with Seq2Seq learning have significantly improved Nash-Sutcliffe efficiency and error metrics [62]. Applications include the downscaling of satellite soil moisture data and geotechnical engineering, where Bi-LSTM consistently outperforms traditional models in complex scenarios [63–65]. The model consists of five stacked Bi-LSTM layers, with hidden units progressively reduced from 128 to 32 to capture both high- and low-level temporal features. Each Bi-LSTM layer is followed by Batch Normalization and Dropout (0.3) to improve generalization and training stability. The final output is produced through a Dense layer with a single unit.

5) Artificial neural networks (ANNs) are versatile tools capable of modeling nonlinear dynamics and noisy data, outperforming traditional hydrological models such as ARMAX and SAC-SMA [66]. They have proven successful in precipitation forecasting and soil erosion studies, where their integration with advanced methods, such as bootstrap aggregation and genetic algorithms, improves accuracy and spatial precision [67]. Robust training algorithms, such as the Levenberg-Marquardt method, further optimize the performance of ANNs [68].

6) Linear regression models with regularization, such as Ridge, Lasso, and ElasticNet, are effective against multicollinearity and overfitting. Lasso uses L1 regularization for feature selection, while Ridge uses L2 regularization to stabilize the coefficient estimates. ElasticNet combines these strengths, making it more robust for complex datasets [69]. These models have shown promise in soil and environmental studies as they balance predictive power and computational efficiency.

7) Least absolute shrinkage and selection operator (Lasso) regression represents the fusion of statistical modeling and machine learning. Consequently, the model can anticipate outcomes and understand the relationships and patterns seen in the data. Because of the controllability of the model, predictions can be refined and customized for a particular scenario. Ridge regression employs the L2 Normalization strategy, while Lasso uses the L1 Normalization technique.

8) ElasticNet addresses the shortcomings of Ridge regression and Lasso regression. By using insights from the literature on Lasso and Ridge regression, ElasticNet enhances the model's regularization. Generally, for this study, this state-of-the-art model is more suitable because of its feature selection, robustness, and higher performance over a vast number of dataset variables [69].

4. Methodology

The machine learning algorithms stated above were used to develop prediction models. Surface soil temperature was precited using three climatic factors: atmospheric temperature, relative humidity, and wind speed. Figure 2, which was generated using OpenAI, showcases the prediction schematics of this study.

4.1. Data cleaning

The collected data were subjected to a cleaning process to minimize the noise of the dataset. Noises in atmospheric data are common due to various errors in data gathering and data transfer, among others. Missing values are a significant drawback. Thus, an expert in the environmental sciences conducted a sensitive analysis, and

Figure 2
Schematics of soil temperature prediction

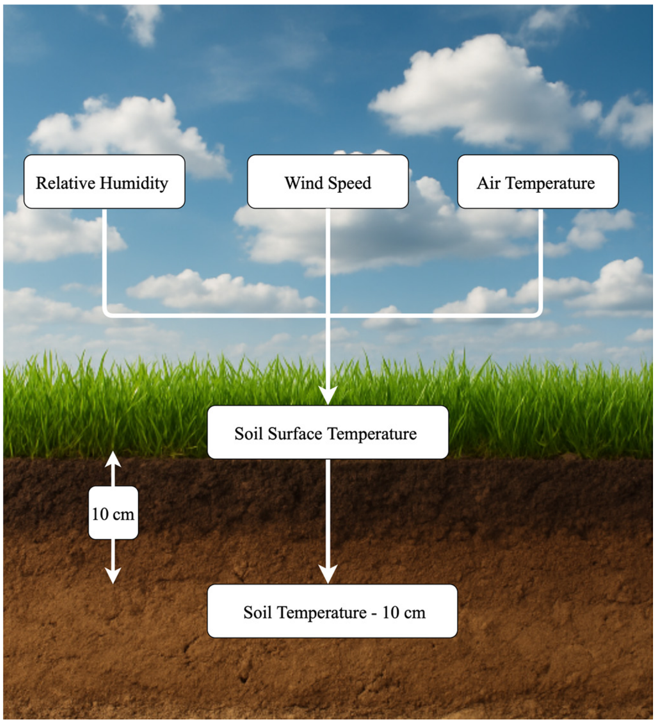
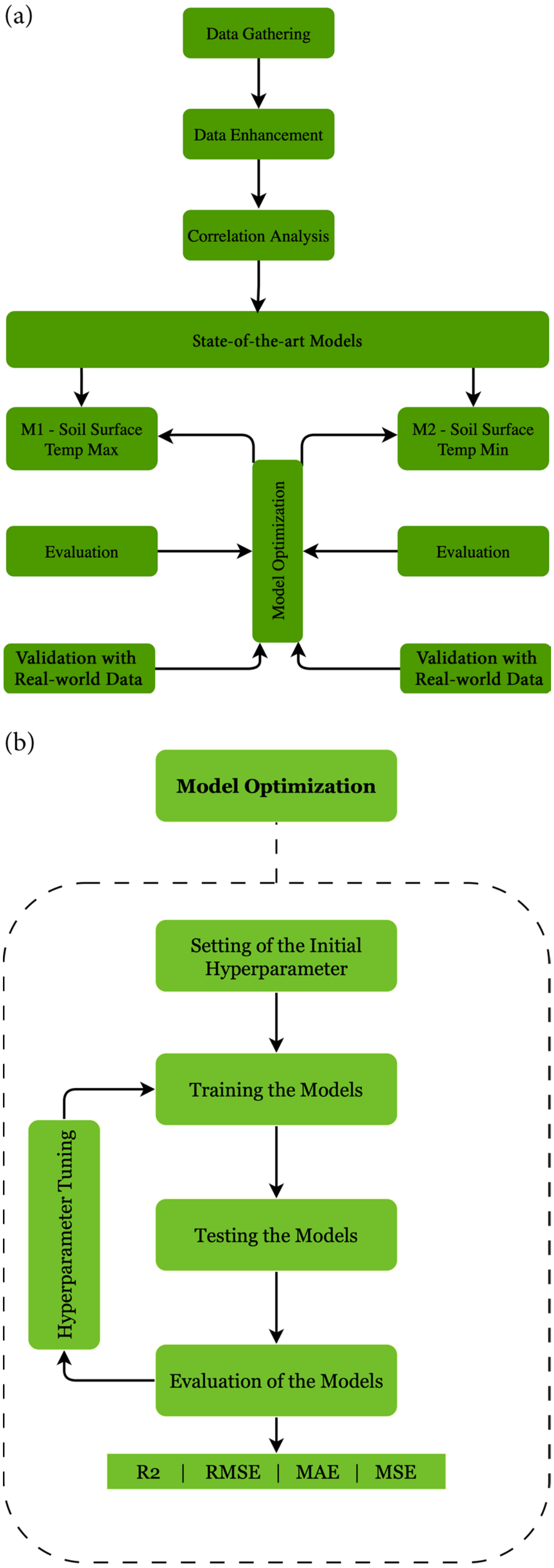


Figure 3
Model-development pipeline and hyperparameter-tuning workflow. (a) Overall methodology; (b) flowchart of the tuning procedure



the dataset was prepared with no missing values (removing the missing data periods). In addition, different mathematical bases such as standard deviations, median absolute deviations, and z-score tests were carried out and reduced the noises of the dataset.

4.2. Mathematical model development

Two mathematical formulations as shown in Equation (1) and Equation (2) were developed using eight machine learning algorithms. Equation (1) was developed to predict the surface soil temperature, whereas Equation (2) showcases the mathematical formulation for the soil temperature at 10 cm level. Equation (2) uses the results from Equation (1) in modeling.

$$T_{surface} = Function(AT_{Max}, AT_{Min}, AT_{Mean}, RH_{Min}, RH_{Mean}, WS) \quad (1)$$

where AT stands for atmospheric temperature, RH stands for relative humidity, and WS is the wind speed.

$$T_{@10cmdepth} = Function(AT_{Max}, AT_{Min}, AT_{Mean}, RH_{Min}, RH_{Mean}, WS, T_{surface}) \quad (2)$$

4.3. Machine learning model development

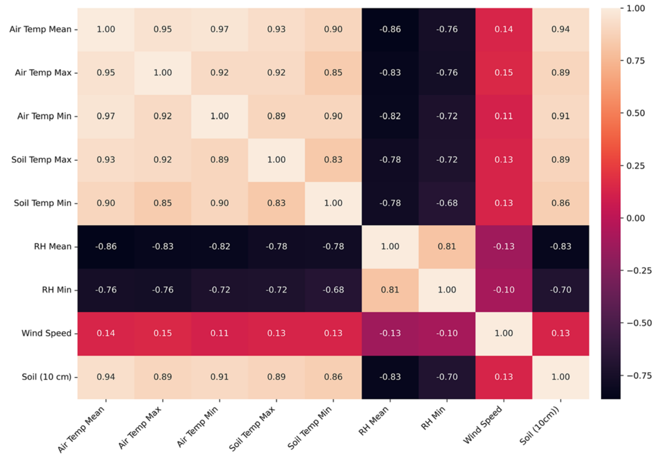
The development and optimization processes are illustrated in Figure 3(a), showcasing the overall machine learning methodology. To mitigate issues of overfitting and underfitting, a systematic approach integrating predictive data evaluation and hyperparameter tuning was employed. Specifically, grid search and k-fold cross-validation techniques were implemented to enhance model efficiency and performance.

Hyperparameter tuning was performed through an iterative process aimed at optimizing model performance. Initially, the models were trained using default hyperparameter settings, and their performance was evaluated based on predefined metrics. In cases of identified errors or suboptimal outcomes, subsequent rounds of hyperparameter tuning were conducted, involving adjustment and re-evaluation of parameters. This iterative cycle continued until the models achieved optimal performance, thereby ensuring robust and accurate predictions. The hyperparameter tuning process is detailed in Figure 3(b).

The best algorithm selected was used to predict the soil temperature levels at a depth of 10 cm. The prediction was carried out in two methods as shown below.

1) Measured data were fed into Equation (2) for the prediction.

Figure 4
Correlation matrix among parameters



2) Measured data for atmospheric parameters were fed with predicted soil surface temperatures into Equation (2).

4.4. Evaluation of prediction models

Widely used evaluation matrices (Regression coefficient (R^2), root mean square error (RMSE), mean absolute error (MAE), and mean square error (MSE)) were used in evaluating and identifying the best prediction models. More details of this evaluation matrices can be found in article by Mampitiya et al. [32].

5. Results and Discussion

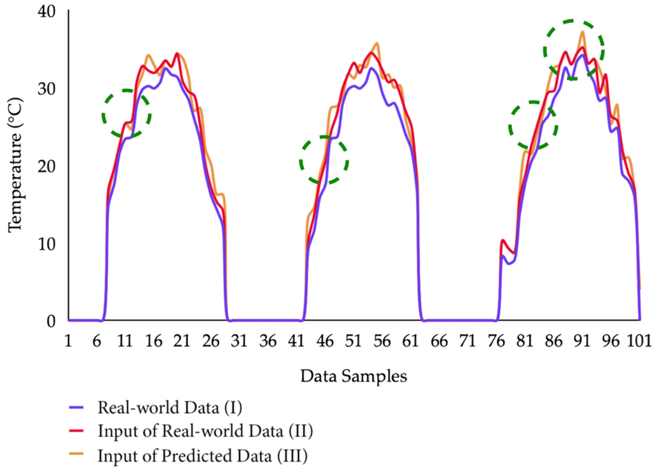
Figure 4 presents the correlation matrix obtained for parameters used in the development of machine learning models. Even though it is linear, the heat map showcases some interesting correlations among considered parameters, including positive correlations from air to soil temperatures. As expected, when atmospheric temperature rises, soil temperature also rises. However, relative humidity shows a negative correlation with both air and soil temperatures. Nevertheless, wind speed has a weak relationship with all other parameters.

Table 1 presents the performance of each machine learning algorithm in predicting soil temperatures at both surface and 10 cm depth. The Bi-LSTM algorithm outperformed other algorithms in

Table 1
Performance evaluation matrices

	Model	R^2	MSE	MAE	RMSE
Maximum soil temperature	XGBoost	0.7969	46.1784	5.4932	6.7954
	CatBoost	0.8023	44.9466	5.5737	6.7042
	LSTM	0.7498	56.875	6.5097	7.5415
	ANN	0.4764	119.0591	9.0597	10.9114
	Bi-LSTM	***0.8497***	***34.1754***	***5.0605***	***5.8459***
	Ridge	0.7234	62.8953	7.1696	7.9306
	Lasso	0.7286	61.7029	7.2562	7.8551
	ElasticNet	0.7223	63.1335	7.2212	7.9456
	Model	**R^2**	**MSE**	**MAE**	**RMSE**
Minimum soil temperature	XGBoost	0.9759	3.2626	1.309	1.8062
	CatBoost	0.9812	2.5457	1.1624	1.5955
	LSTM	0.9629	5.0298	1.6965	2.2427
	ANN	0.984	2.1744	1.1679	1.4745
	Bi-LSTM	***0.9862***	***1.8679***	***1.0101***	***1.3667***
	Ridge	0.9835	2.2354	1.1532	1.4951
	Lasso	0.9809	2.584	1.2748	1.6074
	ElasticNet	0.9816	2.4955	1.2469	1.5797
	Model	**R^2**	**MSE**	**MAE**	**RMSE**
At 10 cm depth	XGBoost	0.9263	12.463	2.0772	3.5303
	CatBoost	0.9271	12.3308	2.1383	3.5115
	LSTM	0.935	10.2657	2.0806	3.204
	ANN	0.8821	17.7499	3.0853	4.213
	Bi-LSTM	***0.9427***	***9.6938***	***1.9424***	***3.113***
	Ridge	0.9047	16.1331	3.0866	4.0166
	Lasso	0.9057	15.9594	3.0569	3.9949
	ElasticNet	0.9049	16.0862	3.0777	4.0107

Figure 5
Comparison of soil temperature with real-world data



predicting soil temperatures for maximum, minimum, and 10 cm depth. The algorithm showcased the highest coefficient of determination, with minimum MSE and minimum absolute errors. However, all other algorithms performed well with high (0.7 or above) R^2 values with relatively low MSE and MAE values.

The models were trained and optimized on a computer equipped with an M2 Neural Engine processor (8 cores) and 16GB RAM. Due to effective optimization and the low complexity of the dataset, training times were consistently below 3 minutes, with inference times under 30 seconds on equivalent hardware. As these models are not intended for deployment on low-power devices and the time differences were negligible, training and inference times were not considered primary evaluation criteria.

Therefore, the Bi-LSTM algorithm was used to predict soil temperature levels at 10 cm depth. Figure 5 presents the comparative analysis of the soil temperature levels at 10 cm depth with real-world data (measured). The purple curve serves as the baseline for predictions. The red curve is the predicted soil temperature at 10 cm depth using the measured parameters. The marked points in the curves demonstrate the models' capability of following the trends of thermal variability.

The actual temperature data were compared with reconstructed or predicted versions. The curves closely follow the measured soil temperature curve, but it has some slightly higher temperature peaks. The red curve (Input of real-world data II) shows the overall shape of the purple curve. The orange curve (Input of predicted data III) mimics the overall shape of the purple curve, with slightly sharper peaks and transitions, indicating that it may be a reconstruction of the model or prediction based on learned patterns from real-world data. Nevertheless, the close alignment of the orange curve with the blue curve suggests that the predicted input (III) is reasonably accurate in capturing the general pattern and magnitude of temperature variation. Minor deviations may occur due to the limitations of the predictive model. Therefore, the prediction models developed based on Bi-LSTM algorithm can be considered for real world implementation.

6. Conclusion

In this study, climatic factors were directly employed in the developed models to predict the minimum and maximum temperatures of the soil surface. Based on climatic data combined with the minimum and maximum surface temperatures, the soil temperature at 10 cm depth was predicted using Bi-LSTM. Bi-LSTM was able to precisely identify the climatic patterns and soil temperature patterns with a higher R^2 score above 0.9. Furthermore, to validate the process, different plots were used to identify whether the models were capable of functioning well with the real-world data. With the following of proper techniques on data handling, machine learning model usage, and validation, this research work concludes that with the existing data, the soil temperature in Bustan, Uzbekistan, is predictable. Therefore, this research work can be utilized to increase crop yields, maintain the irrigation systems properly, and identify the variations that are possible with climatic variations. This research work successfully illustrates that the real-world implication of this model can work precisely with real-time data.

7. Limitations and Future Work

This research offers several advantages that warrant further exploration. Notably, users are not required to measure soil temperatures directly with instruments; instead, they can make predictions utilizing climatic data. However, a significant limitation is the restricted availability of data necessary to broaden the range of input parameters, as well as the lack of comprehensive soil type information, both of which constrain the model's accuracy and applicability. Future development of this research includes the potential to create a real-time soil thermal modeling system integrated with smart sensors. The results will be accessible through a customer interface (web app). Additionally, a forecasting model for early warning applications can also be developed. These advancements will significantly support farmers and policymakers in making informed decisions, thereby contributing to long-term sustainability.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data are available on request from the corresponding author upon reasonable request.

Use of Generative AI and AI-Assisted Tools

Artificial intelligence tools were used solely to enhance the clarity and visual quality of Figure 2.

Author Contribution Statement

Lakindu Mampitiya: Methodology, Software, Formal analysis, Investigation, Writing – original draft, Visualization. **Namal Rathnayake:** Validation, Investigation, Writing – original draft, Writing – review & editing. **Kenjabek Rozumbetov:** Resources, Data curation. **Valery Erkudov:** Resources, Data curation. **Mirzohid Koriyev:** Resources, Data curation. **Komali Kantamaneni:** Writing – review & editing. **Upaka Rathnayake:** Conceptualization, Validation, Writing – review & editing, Supervision, Project administration.

References

[1] Kaur, R., Meier, C. J., McGraw, E. A., Hillyer, J. F., & Bordenstein, S. R. (2024). The mechanism of cytoplasmic incompatibility is conserved in Wolbachia-infected Aedes aegypti mosquitoes

deployed for arbovirus control. *PLoS Biology*, *22*(3), e3002573. https://doi.org/10.1371/journal.pbio.3002573
[2] Robinson, J. M., Liddicoat, C., Muñoz-Rojas, M., & Breed, M. F. (2024). Restoring soil biodiversity. *Current Biology*, *34*(9), R393–R398. https://doi.org/10.1016/j.cub.2024.02.035
[3] Bhaduri, D., Sihi, D., Bhowmik, A., Verma, B. C., Munda, S., & Dari, B. (2022). A review on effective soil health bio-indicators for ecosystem restoration and sustainability. *Frontiers in Microbiology*, *13*, 938481. https://doi.org/10.3389/fmicb.2022.938481
[4] Dincă, L. C., Grenni, P., Onet, C., & Onet, A. (2022). Fertilization and soil microbial community: A review. *Applied Sciences*, *12*(3), 1198. https://doi.org/10.3390/app12031198
[5] Cao, Z., Xue, Y., Nayak, H. P., Lettenmaier, D. P., Frankenberg, C., Köhler, P., & Li, Z. (2024). Understanding terrestrial water and carbon cycles and their interactions using integrated SMAP soil moisture and OCO-2 SIF observations and land surface models. *Journal of Geophysical Research: Atmospheres*, *129*(18), e2024JD041077. https://doi.org/10.1029/2024JD041077
[6] Xiong, M., & Leng, G. (2024). Global soil water erosion responses to climate and land use changes. *Catena*, *241*, 108043. https://doi.org/10.1016/j.catena.2024.108043
[7] Bereswill, S., Gatz-Miller, H., Su, D., Tötzke, C., Kardjilov, N., Oswald, S. E., & Mayer, K. U. (2023). Coupling non-invasive imaging and reactive transport modeling to investigate water and oxygen dynamics in the root zone. *Vadose Zone Journal*, *22*(5), e20268. https://doi.org/10.1002/vzj2.20268
[8] Lynch, J. P., Mooney, S. J., Strock, C. F., & Schneider, H. M. (2022). Future roots for future soils. *Plant, Cell & Environment*, *45*(3), 620–636. https://doi.org/10.1111/pce.14213
[9] de Deyn, G. B., & Kooistra, L. (2021). The role of soils in habitat creation, maintenance and restoration. *Philosophical Transactions of the Royal Society B*, *376*(1834), 20200170. https://doi.org/10.1098/rstb.2020.0170
[10] Sarkar, S., Kumar, R., Kumar, A., Kumar, U., Singh, D. K., Mondal, S., ..., & Yadav, S. K. (2022). Role of soil microbes to assess soil health. In B. Giri, R. Kapoor, Q.-S. Wu, & A. Varma (Eds.), *Structure and functions of pedosphere* (pp. 339–363). Springer. https://doi.org/10.1007/978-981-16-8770-9_14
[11] Bhatti, A. A., Haq, S., & Bhat, R. A. (2017). Actinomycetes benefaction role in soil and plant health. *Microbial Pathogenesis*, *111*, 458–467. https://doi.org/10.1016/j.micpath.2017.09.036
[12] Davis, A. G., Huggins, D. R., & Reganold, J. P. (2023). Linking soil health and ecological resilience to achieve agricultural sustainability. *Frontiers in Ecology and the Environment*, *21*(3), 131–139. https://doi.org/10.1002/fee.2594
[13] Lehmann, J., Bossio, D. A., Kögel-Knabner, I., & Rillig, M. C. (2020). The concept and future prospects of soil health. *Nature Reviews Earth & Environment*, *1*(10), 544–553. https://doi.org/10.1038/s43017-020-0080-8
[14] Luo, Z., Viscarra-Rossel, R. A., & Qian, T. (2021). Similar importance of edaphic and climatic factors for controlling soil organic carbon stocks of the world. *Biogeosciences*, *18*(6), 2063–2073. https://doi.org/10.5194/bg-18-2063-2021
[15] Mani, S., Merino, A., Garcia-Oliva, F., Riotte, J., & Sukumar, R. (2018). Soil properties and organic matter quality in relation to climate and vegetation in southern Indian tropical ecosystems. *Soil Research*, *56*(1), 80–90. https://doi.org/10.1071/SR16262
[16] Galluzzi, G., Plaza, C., Priori, S., Giannetta, B., & Zaccone, C. (2024). Soil organic matter dynamics and stability: Climate vs. time. *Science of the Total Environment*, *929*, 172441. https://doi.org/10.1016/j.scitotenv.2024.172441
[17] Cholet, C., Houle, D., Sylvain, J.-D., Doyon, F., & Maheu, A. (2022). Climate change increases the severity and duration of soil water stress in the temperate forest of Eastern North America. *Frontiers in Forests and Global Change*, *5*, 879382. https://doi.org/10.3389/ffgc.2022.879382
[18] Peplau, T., Poeplau, C., Gregorich, E., & Schroeder, J. (2023). Deforestation for agriculture leads to soil warming and enhanced litter decomposition in subarctic soils. *Biogeosciences*, *20*(5), 1063–1074. https://doi.org/10.5194/bg-20-1063-2023
[19] Jungqvist, G., Oni, S. K., Teutschbein, C., & Futter, M. N. (2014). Effect of climate change on soil temperature in Swedish boreal forests. *PloS One*, *9*(4), e93957. https://doi.org/10.1371/journal.pone.0093957
[20] Pepper, I. L., & Brusseau, M. L. (2019). Physical-chemical characteristics of soils and the subsurface. In M. L. Brusseau, I. L. Pepper, & C. P. Gerba (Eds.), *Environmental and pollution science* (9–22). Academic Press. https://doi.org/10.1016/B978-0-12-814719-1.00002-1
[21] Cruz-Paredes, C., Tájmel, D., & Rousk, J. (2023). Variation in temperature dependences across Europe reveals the climate sensitivity of soil microbial decomposers. *Applied and Environmental Microbiology*, *89*(5), e02090-22. https://doi.org/10.1128/aem.02090-22
[22] Daunoras, J., Kačergius, A., & Gudiukaitė, R. (2024). Role of soil microbiota enzymes in soil health and activity changes depending on climate change and the type of soil ecosystem. *Biology*, *13*(2), 85. https://doi.org/10.3390/biology13020085
[23] Fang, C., Verbrigghe, N., Sigurdsson, B. D., Ostonen, I., Leblans, N. I., Marañón-Jiménez, S., ..., & Janssens, I. A. (2023). Decadal soil warming decreased vascular plant above and belowground production in a subarctic grassland by inducing nitrogen limitation. *New Phytologist*, *240*(2), 565–576. https://doi.org/10.1111/nph.19177
[24] Kumar, L., Chhogyel, N., Gopalakrishnan, T., Hasan, M. K., Jayasinghe, S. L., Kariyawasam, C. S., ..., & Ratnayake, S. (2022). Climate change and future of agri-food production. In R. Bhat (Ed.), *Future foods: Global trends, opportunities, and sustainability challenges* (pp. 49–79). Academic Press. https://doi.org/10.1016/B978-0-323-91001-9.00009-8
[25] Habib-ur-Rahman, M., Ahmad, A., Raza, A., Hasnain, M. U., Alharby, H. F., Alzahrani, Y. M., ..., & El Sabagh, A. (2022). Impact of climate change on agricultural production; Issues, challenges, and opportunities in Asia. *Frontiers in Plant Science*, *13*, 925548. https://doi.org/10.3389/fpls.2022.925548
[26] Muluneh, M. G. (2021). Impact of climate change on biodiversity and food security: A global perspective—A review article. *Agriculture & Food Security*, *10*(1), 36. https://doi.org/10.1186/s40066-021-00318-5
[27] Mirón, I. J., Linares, C., & Díaz, J. (2023). The influence of climate change on food production and food safety. *Environmental Research*, *216*, 114674. https://doi.org/10.1016/j.envres.2022.114674
[28] Dasgupta, S., & Robinson, E. J. (2022). Attributing changes in food insecurity to a changing climate. *Scientific Reports*, *12*(1), 4709. https://doi.org/10.1038/s41598-022-08696-x
[29] Wang, W., Wang, W., Wang, P., Wang, X., Wang, L., Wang, C., ..., & Huo, Z. (2023). Impact of straw return on soil temperature and water during the freeze-thaw period. *Agricultural Water Management*, *282*, 108292. https://doi.org/10.1016/j.agwat.2023.108292
[30] Dai, Z., Hu, J., Fan, J., Fu, W., Wang, H., & Hao, M. (2021). No-tillage with mulching improves maize yield in dryland

farming through regulating soil temperature, water and nitrate-N. *Agriculture, Ecosystems & Environment*, *309*, 107288. https://doi.org/10.1016/j.agee.2020.107288

[31] Bibi, F., & Rahman, A. (2023). An overview of climate change impacts on agriculture and their mitigation strategies. *Agriculture*, *13*(8), 1508. https://doi.org/10.3390/agriculture13081508

[32] Mampitiya, L., Rozumbetov, K., Rathnayake, N., Erkudov, V., Esimbetov, A., Arachchi, S., ..., & Rathnayake, U. (2024). Artificial intelligence to predict soil temperatures by development of novel model. *Scientific Reports*, *14*(1), 9889. https://doi.org/10.1038/s41598-024-60549-x

[33] Kholmatjanov, B. M., Petrov, Y. V., Khujanazarov, T., Sulaymonova, N. N., Abdikulov, F. I., & Tanaka, K. (2020). Analysis of temperature change in Uzbekistan and the regional atmospheric circulation of Middle Asia during 1961–2016. *Climate*, *8*(9), 101. https://doi.org/10.3390/cli8090101

[34] Tahat, M. M., Alananbeh, K. M., Othman, Y. A., & Leskovar, D. I. (2020). Soil health and sustainable agriculture. *Sustainability*, *12*(12), 4859. https://doi.org/10.3390/su12124859

[35] Montanarella, L., & Panagos, P. (2018). Soil data needs for sustainable agriculture. In H. Ginzky, E. Dooley, I. L. Heuser, E. Kasimbazi, T. Markus, & T. Qin (Eds.), *International Yearbook of Soil Law and Policy 2017* (pp. 151–166). Springer International Publishing. https://doi.org/10.1007/978-3-319-68885-5_9

[36] Koriyev, M. R., Fonseka, P. U., Umurzakova, U. N., Rozumbetov, K. U., Arachchi, S., Erkudov, V. O., ..., & Rathnayake, U. (2024). Soil salinity status in namangan region, uzbekistan. *Suranaree Journal of Science and Technology*, *31*(5), 010335. https://doi.org/10.55766/sujst-2024-05-e05652

[37] Farhangmehr, V., Imanian, H., Mohammadian, A., Cobo, J. H., Shirkhani, H., & Payeur, P. (2025). A spatiotemporal CNN-LSTM deep learning model for predicting soil temperature in diverse large-scale regional climates. *Science of The Total Environment*, *968*, 178901. https://doi.org/10.1016/j.scitotenv.2025.178901

[38] Imanian, H., Mohammadian, A., Farhangmehr, V., Payeur, P., Goodarzi, D., Hiedra Cobo, J., & Shirkhani, H. (2024). A comparative analysis of deep learning models for soil temperature prediction in cold climates. *Theoretical and Applied Climatology*, *155*(4), 2571–2587. https://doi.org/10.1007/s00704-023-04781-x

[39] Malik, A., Tikhamarine, Y., Sihag, P., Shahid, S., Jamei, M., & Karbasi, M. (2022). Predicting daily soil temperature at multiple depths using hybrid machine learning models for a semi-arid region in Punjab, India. *Environmental Science and Pollution Research*, *29*(47), 71270–71289. https://doi.org/10.1007/s11356-022-20837-3

[40] Bayatvarkeshi, M., Bhagat, S. K., Mohammadi, K., Kisi, O., Farahani, M., Hasani, A., ..., & Yaseen, Z. M. (2021). Modeling soil temperature using air temperature features in diverse climatic conditions with complementary machine learning models. *Computers and Electronics in Agriculture*, *185*, 106158. https://doi.org/10.1016/j.compag.2021.106158

[41] Abimbola, O. P., Meyer, G. E., Mittelstet, A. R., Rudnick, D. R., & Franz, T. E. (2021). Knowledge-guided machine learning for improving daily soil temperature prediction across the United States. *Vadose Zone Journal*, *20*(5), e20151. https://doi.org/10.1002/vzj2.20151

[42] Alizamir, M., Kisi, O., Ahmed, A. N., Mert, C., Fai, C. M., Kim, S., ..., & El-Shafie, A. (2020). Advanced machine learning model for better prediction accuracy of soil temperature at different depths. *PLoS One*, *15*(4), e0231055. https://doi.org/10.1371/journal.pone.0231055

[43] Feng, Y., Cui, N., Hao, W., Gao, L., & Gong, D. (2019). Estimation of soil temperature from meteorological data using different machine learning models. *Geoderma*, *338*, 67–77. https://doi.org/10.1016/j.geoderma.2018.11.044

[44] Rakhmatullaev, S., Huneau, F., le Coustumer, P., Motelica-Heino, M., & Bakiev, M. (2010). Facts and perspectives of water reservoirs in Central Asia: A special focus on Uzbekistan. *Water*, *2*(2), 307–320. https://www.mdpi.com/2073-4441/2/2/307#

[45] Abdullaev, I., Rakhmatullaev, S., Platonov, A., & Sorokin, D. (2012). Improving water governance in Central Asia through application of data management tools. *International Journal of Environmental Studies*, *69*(1), 151–168. https://doi.org/10.1080/00207233.2011.641243

[46] Al-Kaisi, M. M., Lal, R., Olson, K. R., & Lowery, B. (2017). Fundamentals and functions of soil environment. In M. M. Al-Kaisi & B. Lowery (Eds.), *Soil health and intensification of agroecosystems* (pp. 1–23). Academic Press. https://doi.org/10.1016/B978-0-12-805317-1.00001-4

[47] Howe, J. A., & Smith, A. P. (2021). The soil habitat. In T. J. Gentry, J. J. Fuhrmann, & D. A. Zuberer (Eds.), *Principles and applications of soil microbiology* (pp. 23–55). Elsevier. https://doi.org/10.1016/B978-0-12-820202-9.00002-2

[48] Makumbura, R. K., Mampitiya, L., Rathnayake, N., Meddage, D. P. P., Henna, S., Dang, T. L., ..., & Rathnayake, U. (2024). Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature. *Results in Engineering*, *23*, 102831. https://doi.org/10.1016/j.rineng.2024.102831

[49] Mustapha, I. B., & Saeed, F. (2016). Bioactive molecule prediction using extreme gradient boosting. *Molecules*, *21*(8), 983. https://doi.org/10.3390/molecules21080983

[50] Yoon, S.-H., Jeon, J.-H., Cho, S.-B., Nacpil, E. J. C., Jeon, I., Choi, J.-B., & Kim, H. (2023). Extreme gradient boosting to predict atomic layer deposition for platinum nano-film coating. *Langmuir*, *39*(14), 4984–4992. https://doi.org/10.1021/acs.langmuir.2c03465

[51] Georganos, S., Grippa, T., Vanhuysse, S., Lennert, M., Shimoni, M., & Wolff, E. (2018). Very high resolution object-based land use–land cover urban classification using extreme gradient boosting. *IEEE Geoscience and Remote Sensing Letters*, *15*(4), 607–611. https://doi.org/10.1109/LGRS.2018.2803259

[52] Chen, B., Zheng, H., Luo, G., Chen, C., Bao, A., Liu, T., & Chen, X. (2022). Adaptive estimation of multi-regional soil salinization using extreme gradient boosting with Bayesian TPE optimization. *International Journal of Remote Sensing*, *43*(3), 778–811. https://doi.org/10.1080/01431161.2021.2009589

[53] Ye, Z., Sheng, Z., Liu, X., Ma, Y., Wang, R., Ding, S., ..., & Wang, Q. (2021). Using machine learning algorithms based on GF-6 and google earth engine to predict and map the spatial distribution of soil organic matter content. *Sustainability*, *13*(24), 14055. https://doi.org/10.3390/su132414055

[54] Samat, A., Li, E., Wang, W., Liu, S., Lin, C., & Abuduwaili, J. (2020). Meta-XGBoost for hyperspectral image classification using extended MSER-guided morphological profiles. *Remote Sensing*, *12*(12), 1973. https://doi.org/10.3390/rs12121973

[55] Anand, N., Singh, J. P., Gourisaria, M. K., Bandyopadhyay, A., Singh, V., & Patra, S. S. (2024). Oil spill classification: A machine learning approach. In *2024 2nd International Conference on Sustainable Computing and Smart Systems*, 834–840. https://doi.org/10.1109/ICSCSS60660.2024.10625513

[56] Zhou, Y., Wu, Z., Jiang, M., Xu, H., Yan, D., Wang, H., ..., & Zhang, X. (2024). Real-time prediction and ponding process early warning method at urban flood points based on different deep learning methods. *Journal of Flood Risk Management*, *17*(1), e12964. https://doi.org/10.1111/jfr3.12964

[57] Sahaar, S. A., & Niemann, J. D. (2024). Estimating rootzone soil moisture by fusing multiple remote sensing products with machine learning. *Remote Sensing*, *16*(19), 3699. https://doi.org/10.3390/rs16193699

[58] Kratzert, F., Klotz, D., Brenner, C., Schulz, K., & Herrnegger, M. (2018). Rainfall–runoff modelling using long short-term memory (LSTM) networks. *Hydrology and Earth System Sciences*, *22*(11), 6005–6022. https://doi.org/10.5194/hess-22-6005-2018

[59] Li, C., & Fang, H. (2021). Assessment of climate change impacts on the streamflow for the Mun River in the Mekong Basin, Southeast Asia: Using SWAT model. *Catena*, *201*, 105199. https://doi.org/10.1016/j.catena.2021.105199

[60] Waqas, M., & Humphries, U. W. (2024). A critical review of RNN and LSTM variants in hydrological time series predictions. *MethodsX*, *13*, 102946. https://doi.org/10.1016/j.mex.2024.102946

[61] Cai, H., Liu, S., Shi, H., Zhou, Z., Jiang, S., & Babovic, V. (2022). Toward improved lumped groundwater level predictions at catchment scale: Mutual integration of water balance mechanism and deep learning method. *Journal of Hydrology*, *613*, 128495. https://doi.org/10.1016/j.jhydrol.2022.128495

[62] Li, F., Ma, G., Chen, S., & Huang, W. (2021). An ensemble modeling approach to forecast daily reservoir inflow using bidirectional long-and short-term memory (Bi-LSTM), variational mode decomposition (VMD), and energy entropy method. *Water Resources Management*, *35*(9), 2941–2963. https://doi.org/10.1007/s11269-021-02879-3

[63] De, S., Mukherjee, P., & Roy, A. H. (2025). TasteNet: A novel deep learning approach for EEG-based basic taste perception recognition using CEEMDAN domain entropy features. *Journal of Neuroscience Methods*, *419*, 110463. https://doi.org/10.1016/j.jneumeth.2025.110463

[64] Koné, B. A. T., Grati, R., Bouaziz, B., & Boukadi, K. (2023). A new long short-term memory based approach for soil moisture prediction. *Journal of Ambient Intelligence and Smart Environments*, *15*(3), 255–268. https://doi.org/10.3233/AIS-230035

[65] Yu, Y., Yao, Y., Liu, Z., An, Z., Chen, B., Chen, L., & Chen, R. (2023). A Bi-LSTM approach for modelling movement uncertainty of crowdsourced human trajectories under complex urban environments. *International Journal of Applied Earth Observation and Geoinformation*, *122*, 103412. https://doi.org/10.1016/j.jag.2023.103412

[66] da Silva, D. D., Uliana, E. M., Moreira, M. C., & Pereira, S. B. (2016). Applying the SAC-SMA hydrologic model to a tropical watershed in Brazil. In *2016 ASABE Annual International Meeting*. https://doi.org/10.13031/aim.20162460855

[67] Langella, G., Basile, A., Bonfante, A., & Terribile, F. (2010). High-resolution space–time rainfall analysis using integrated ANN inference systems. *Journal of Hydrology*, *387*(3–4), 328–342. https://doi.org/10.1016/j.jhydrol.2010.04.027

[68] Ly, H.-B., Nguyen, M. H., & Pham, B. T. (2021). Metaheuristic optimization of Levenberg–Marquardt-based artificial neural network using particle swarm optimization for prediction of foamed concrete compressive strength. *Neural Computing and Applications*, *33*(24), 17331–17351. https://doi.org/10.1007/s00521-021-06321-y

[69] Mampitiya, L., Rathnayake, N., Hoshino, Y., & Rathnayake, U. (2024). Forecasting PM10 levels in Sri Lanka: A comparative analysis of machine learning models PM10. *Journal of Hazardous Materials Advances*, *13*, 100395. https://doi.org/10.1016/j.hazadv.2023.100395

How to Cite: Mampitiya, L., Rathnayake, N., Rozumbetov, K., Erkudov, V., Koriyev, M., Kantamaneni, K., & Rathnayake, U. (2026). AI-Powered Soil Temperature Modeling for Sustainable Agriculture in Arid Regions: A Case Study of Bustan, Uzbekistan. *Journal of Data Science and Intelligent Systems*. https://doi.org/10.47852/bonviewJDSIS62026463