ELECTRICAL ENGINEERING AND ENERGY

OPEN ACCESS

Air Temperature Forecasting Using a Long Short-Term Memory Model: A Case Study in Beykoz, Istanbul

Gizem Temelcan Ergenecoşar [1], İnal Begüm Turna Demirel [2]

[1] Department of Software Engineering, Faculty of Engineering and Architecture, Beykoz University, Beykoz-Istanbul, Türkiye
[2] Department of Computer Engineering, Faculty of Engineering and Architecture, Beykoz University, Beykoz-Istanbul, Türkiye

Submitted: 19 June 2026
Accepted: 07 August 2026
Online First: 14 August 2026

Corresponding author
İnal Begüm Turna Demirel
inalbegumturna@beykoz.edu.tr

DOI: 10.64470/elene.2026.39

Abstract: Human activities contribute to the production of greenhouse gases, which in turn lead to a rise in temperature and an increase in carbon footprint. With increasing frequency of extreme temperature events due to urban heat island effects and climate change, short-term high-resolution temperature forecasting is critical for localized energy management and urban climate adaptation. In this study, real-world atmospheric data, including as wind direction and speed, temperature, humidity, and precipitation amount are collected via the PCE-FWS 20N weather station in Beykoz, Istanbul. Applying machine learning models to actual atmospheric data demonstrates the potential for enhanced environmental predictions and promotes informed decision-making to ensure a sustainable future. A Long Short-Term Memory (LSTM) machine learning model with optimization methods ADAM (Adaptive Moment Estimation), NADAM (Nesterov-accelerated Adaptive Moment Estimation), and RMSprop (Root Mean Square Propagation) are applied independently for different epoch sizes. A comparative evaluation of the selected optimization algorithms was performed using RMSE and R^2 metrics to identify the reliable configuration for the analyzed dataset.

Keywords Air temperature forecasting, Machine learning, Long short-term memory, Optimization methods

1. Introduction

Member countries of the United Nations issued a call to action that included 17 primary objectives in 2016. The Sustainable Development Goals (SDGs) encompass a comprehensive set of objectives addressing social, cultural, and ecological challenges. This collective endeavor is intended to be accomplished by the conclusion of 2030. The 13th objective, climate action, holds significant importance in this study, as well as in numerous other subjects. Santamouries et al. (2015) presented a paper stating that a 1℃ rise in air temperature led to an increase in electricity demand for buildings ranging from 0.5 to 8.5. According to study, the consumption of energy in buildings makes up approximately 31.4 of the total world energy use. Hence, accurate temperature forecasting is crucial in mitigating climate change and constructing cities that are resilient to climatic variations. Furthermore, the accurate calculation of temperature is the fundamental basis for the sustainable use of water resources and drought assessment. Since human activities lead to the formation of greenhouse gases, which cause more solar radiation to be transmitted to the earth's surface and therefore to an increase in temperature, it increases the carbon footprint as a result. At this point, making forward-looking atmosphere forecasts gains importance in terms of sustainability. Therefore, accurate temperature forecast plays a crucial role in sustainability and climate change assessment, as well as in climate change mitigation and resilient city construction.

Recent research has utilized modern technical advancements to make precise forecasting of meteorological data. Specifically, it gains advantages from the utilization of various heuristic-based models such as statistical methods, machine learning (ML), artificial neural networks (ANN), and deep learning (DL). On the other hand, considering the advantages in computational speed, ML models are applied for urban air temperature forecasting even though there are some limitations of machine learning models remain unclear. For example, ML models require large amounts of high-quality data to be effective, as well as no biases, noise, or gaps to affect the forecasts' accuracy. In addition to data quality and availability limitations, overfitting occurs when a model learns the training data highly, including its noise and outliers. It may perform excellently on the training set but poorly on unseen data. Another limitation can be transferability, that is, a model trained on data from one region may not perform well in another due to different environmental factors and urban layouts. Also, traditional ML models may not effectively capture uncertainty in temperature forecasts. Understanding the risks of extreme temperatures, which can be labeled as limited handling of uncertainty, can be crucial for decision-making processes.

Besides implementing ML models, neural networks possess significant benefits in forecasting due to their exceptional non-linear adaptive processing capability. Typically, short-term flow prediction requires the previous several time steps and accounting for dependencies between the time series. Typical neural networks cannot accomplish this task. Utilizing previous events to deduce future events is often achieved through the use of a Recurrent Neural Network (RNN). In the process of training neural networks, the gradient descent algorithm, which minimizes the errors between actual and expected results as much as possible, is used to evaluate the network's performance. However, Back Propagation Through Time (BPTT) method is used in the traditional RNN to train the network by back propagating errors through time. BPTT requires storing the activations of each time step, which can be memory-intensive for long sequences. This is a limitation in the size of the sequence that can be processed by the network. Consequently, a long training time will make the network's residual drop very quickly, which will cause network weights to be renewed slowly. Therefore, a memory block is needed to store memory, and the LSTM model, a special RNN model, is proposed for keeping long-term dependencies and reading any length sequence.

There has been atmospheric weather forecasting and prediction studies conducted in the literature. But before this, it is essential to state the difference between forecast and prediction and which data will be used. Forecasts are typically based on data-driven analysis and modeling techniques, aiming to provide informed estimates of future outcomes. Predictions, on the other hand, may involve a broader range of methods and can vary in terms of their reliability and basis in evidence. Therefore, temporal air temperature forecasting is the main topic of this study. Although forecasting atmospheric temperature can be challenging due to factors such as heterogeneous landscapes and complex atmospheric dynamics, in this study, machine learning has been employed alongside numerical weather simulations to make reliable and accurate temperature forecasts for the selected region, without requiring an excessive amount of physical data. In this study, a temporal analysis has been conducted.

Weather data is collected across the space and time axes, which means it is spatiotemporal. Spatial forecasting deals with predicting temperature variations across different locations at a specific moment, while temporal forecasting is concerned with predicting how temperatures will change over time at a particular location. Both types of forecasting are essential for understanding and planning for weather-related events and impacts.

Articles in the literature can be divided into three categories: spatial forecast, temporal forecast, and spatiotemporal forecast. Wang et al. (2023) separated the studies among 45 articles into two groups according to the usage of ML models and found a result that tree-based models fit spatial forecast and neural network models work for temporal forecast.

Sofi and Oseledets (2024) presented several data-driven approaches for spatiotemporal forecasting and showed how to condition these models for forecasting weather states. They proposed a spatiotemporal forecasting model and found that its predictions are comparable to other ML models without the need for training.

Detecting weather conditions is essential for many reasons such as public safety (by providing early warnings to prevent disasters), agriculture, energy management, transportation, water management. One of the important weather forecasting methods, which is also the subject of this study, is statistical methods, including time series analysis, machine learning, and ensemble forecasting.

There are several Machine Learning (ML) models adopted in weather forecasting including linear regression (Holmstrom et al., 2016; Sharapov, 2022; Apaydin et al., 2022), support vector machines (Dritsas et al., 2022; Nayak and Ghosh, 2013; Deif et al., 2021), decision trees (Dadhich et al., 2021), random forests (Jakaria et al., 2020; Parlak and Yavasoglu, 2023; Kareem et al., 2021), gradient boosting (Era et al., 2023), K-nearest neighbors (KNN) (Moorthy and Parameshwaran, 2022; Findawati et al., 2019), and artificial neural networks (Johnstone and Sulungu, 2021; Shrivastava et al., 2023a, 2023b). Neural networks provide significant benefits in prediction due to their non-linear adaptive processing capacity. Predicting short-term changes in temperature requires the previous several time steps and the relationships between these time series. Traditional neural networks are unable to accomplish this task. Utilizing previous occurrences to predict future occurrences is often accomplished using a Recurrent Neural Network (RNN). Gradient descent method is generally used to evaluate the network performance during training process of neural networks. However, traditional RNN uses back propagation through time algorithm for training. Here, if training process is long, the network residual that needs to be returned will drop rapidly, and thus it slows down the renewal of network weights. The typical RNN cannot reflect long-term memory in actual applications, and consequently a memory block is essential to store memory. Thus, the Long Short-Term Memory (LSTM) model is designed to preserve long-term dependencies and read any length sequence. LSTM networks are a class of RNN. The output from the prior step is utilized as the upcoming step's input in the RNN. These RNNs have the capacity to learn sequence dependency. LSTMs are known to store data for a long time. Unlike RNNs, LSTMs have the skills to remember the information for a long period of time. LSTM has a more complex structure with additional memory cells and gates that allow it to selectively remember or forget information from previous time steps. Compared to other NN models, since LSTMs are the most effective and well-known subgroup of RNNs, they are designed to recognize patterns in sequences of data, particularly numerical time series data. LSTM is first proposed and applied in time series by Hochreiter and Schmidhuber (1997). Then, the LSTM method is used in different fields such as soil moisture (Filipovic et al., 2022; Han et al., 2021, Gao et al., 2021) and temperature prediction (Suleman and Shridevi, 2022; Li et al., 2022; Kumari and Muthulakshmi, 2023; Jaharabi et al., 2023; Katušić et al., 2022), financial market (Fischer and Krauss, 2018; Bao et al., 2017; Huang et al., 2021), petroleum production (Sagheer and Kotb,2019; Kumar et al., 2023), traffic flow (Xiao and Yin, 2019; Shao and Soong, 2016; Fang et al., 2022; Mou et al., 2019). Since weather forecasting is one of the essential applications of time-series analysis and the neural network models (especially the LSTM) are more frequently used than other models, we applied the LSTM model to the temporal forecasting study.

Within the scope of this study; real-world atmospheric data including the attributes wind direction and speed, temperature, relative humidity, and precipitation amount are collected via the PCE-FWS 20N Weather station located in the Beykoz University Undergraduate Programs Laboratory and Ateliers Campus Building, Istanbul, Turkiye. While long-term climate projections provide macro-level trends, urban centers require localized, short-term forecasting to manage immediate operational impacts, such as peak cooling demands during severe heat periods. Therefore, this study focuses on high-resolution, short-term hourly temperature forecasting during the peak summer period (July–September) in Beykoz, Istanbul. This short-term microclimate modeling serves as a functional tool for localized urban energy management and adaptive response during extreme thermal stress periods.

This study evaluates the performance of an LSTM model for short-term hourly air temperature forecasting using data collected from a single weather station located in Beykoz during the summer period, and compares three optimization algorithms which are ADAM, NADAM, RMSprop. The findings should therefore be interpreted within the scope of the analyzed dataset. The statistical analysis method root mean square error is applied between the forecasted and experimental temperature data.

2. Material And Method

This study is carried out in the Beykoz district of Istanbul, Turkey with the support of Beykoz University. Beykoz region has moderate, temperate weather. There is much more precipitation in Beykoz during the winter than during the summer. The average annual temperature during the year varies around 18.6 °C meanwhile average annual rainfall totals for the year is 728 mm.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Avg. Temperature (°C)	5.8	6.4	8.5	12	17	21.7	24.3	24.6	21	16.2	12	7.8
Min. Temperature (°C)	3.4	3.8	5.3	8.3	13.1	17.9	20.7	21.5	18.1	13.7	9.6	5.6
Max. Temperature (°C)	8.2	9.2	11.9	16	20.9	25.5	28.1	28.2	24.3	19	14.7	10.1
Precipitation (mm)	88	75	75	50	38	35	26	24	52	80	78	107
Humidity (%)	79%	77%	74%	73%	71%	68%	67%	68%	69%	75%	77%	78%
Rainy Days (d.)	10	9	8	6	5	4	3	3	5	7	8	10
Sunshine Hours (h)	5.0	5.8	7.4	9.4	10.8	11.8	11.7	10.6	9.1	6.9	6.0	5.0

Figure 1. Beykoz region monthly weather detail (Climate Data, 2024)

The extreme and average values of weather feature is given in Figure 1. It includes atmospheric data between 1991 and 2021. June, July and August were the months when the longest sunshine durations and therefore high average temperatures were observed.

As stated, accurate temperature forecasting has an important role in evaluating and assessing climate change. Therefore, it is important to know temperature changes within wide time range. Figure 2 represents the climate change effect on Beykoz during forty years, covering the time range from 1979 to 2023.

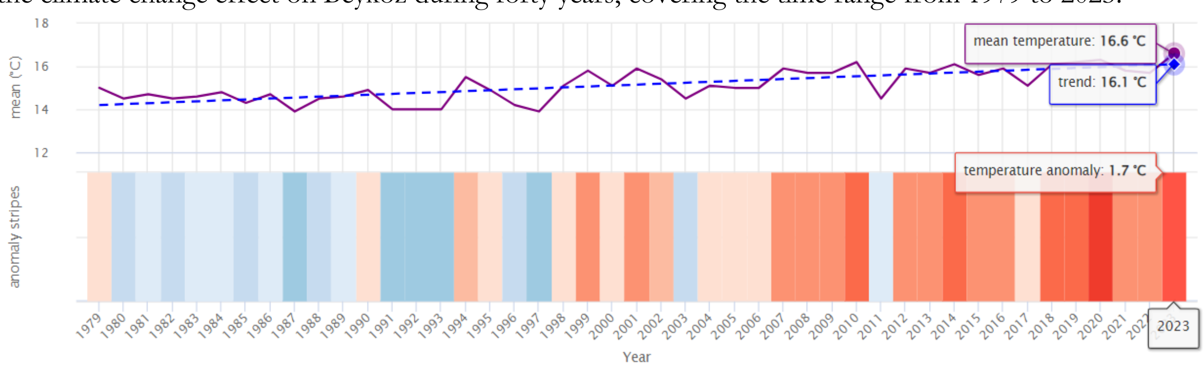

Figure 2. Mean annual temperature, trend and anomaly (Climate Change Beykoz, 2024)

Within this time period, the temperature trend, represented via dashed blue line, is positive which shows that every year the temperature increases due to climate change as if it is 14.2 ℃ in 1979 and 16.1 ℃ in 2023. In the anomaly stripes yearly graph, the blue color for colder and red color for warmer years. In 1979 temperature anomaly is 0.2 ℃ and 1.7 ℃ in 2023. This study focuses on short-term, high-resolution microclimate prediction at times of severe thermal stress, especially in the summer, even though the Figure 2 depicts temperature variance across about 40 years.

2.1. Experiment setup

Data is taken from the PCE-FWS 20N Air Measurement Device which collects atmospheric data such as wind direction and speed, temperature, relative humidity, and precipitation amount. The recorded data is collected through software called Easy Weather.

Table 1. Technical details of Air Measurement Device

Feature	Measurement Range	Resolution	Accuracy
Temperature	-40 … 60 °C	0.1 °C	± 1 °C
Relative Humidity	1 … 99 %	1%	±4% in the range 20…80%, ±6% in other ranges
Precipitation	0 … 9999 mm	0.3 mm (if rain volume <1000 mm) 1 mm (if rain volume > 1000 mm)	±6 %
Wind Speed	0…50 m/s	0.1	±1 m/s (Wind speed <5 m/s) ±10 % (Wind speed >5 m/s)

Technical details of Air Measurement Device are given Table 1 and the area where the device is installed, and the experiment carried out is in Figure 3(a). It is on the roof of the Beykoz University Bachelor's Programs Campus - Labs and Studios building at 41.0924542 latitude and 29.0936903 longitude coordinates. The outdoor unit with sensors is in an open area so that surrounding buildings do not interfere with data collection. Wind speed, direction and precipitation data are collected via wind speed sensor, wind vane and rain collector, respectively. Data is transmitted from the outdoor unit to the LCD screen via antenna.

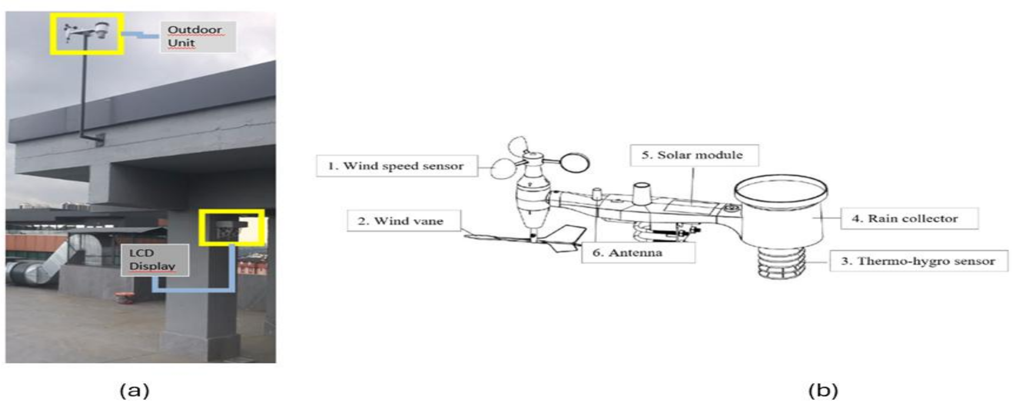

Figure 3. (a) Experiment Setup; outdoor unit and LCD display, (b) PCE-FWS 20N Air Measurement Device (outdoor unit)

The data used for air temperature estimation within the scope of the study is collected by the Thermo-Hygro Sensor shown in Figure 3(b) and recorded via the LCD Display. Within this sensor, there is a thermometer and a hygrometer which measures ambient temperature with ± 1 °C accuracy and the relative humidity of the air with ±4% of accuracy in the range between 20% and 80%, ±6% of accuracy in other ranges.

2.2. Data collection and visualization

The dataset consists of 16 features and 1358 rows that are received from PCE-FWS 20N device within the date range between 7 July and 1 September 2023 spanning 60 minutes. Within the time period the data for temperature is illustrated in Figure 4.

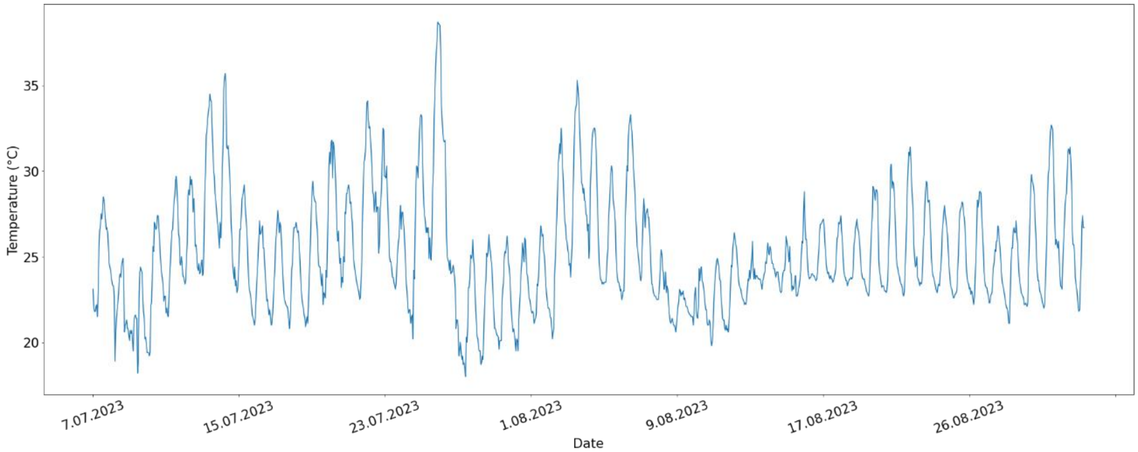

Figure 4. Time period for temperature between July 7-September 1

Outdoor temperature has an important place in this scope of study. The detail for several significant features is stated in Table 2.

Table 2. Extreme collected data points

Feature	Minimum Value	Maximum Value	Mean Value	Standard Deviation
Air Temperature(°C)	18.0	38.7	25.17	3.32
Humidity (%)	25.0	98.0	71.55	15.69
Wind Speed(km/h)	0.0	34.2	7.65	5.50
Total Rainfall(mm)	0.0	31.8	11.95	9.55
Gust(km/h)	0.0	47.9	12.85	8.44
DewPoint(°C)	11.1	24.0	19.23	2.48
WindChill(°C)	16.5	38.7	24.62	3.49

In this part, the correlation matrix of the attributes and the temperature-time graph is illustrated in Figure 5. The correlation matrix shows the relationships between several weather variables. The resulting correlation coefficient within each cell ranges from -1 to 1. The positive correlation, colored in red, variables change similarly. For instance, between Outdoor Temperature and WindChill, with the correlation coefficient of 0.98, when one variable increases, the other increases. On the other hand, temperature and humidity have negative correlation, colored in blue, with the coefficient of -0.75. This result indicates variables changes inversely.

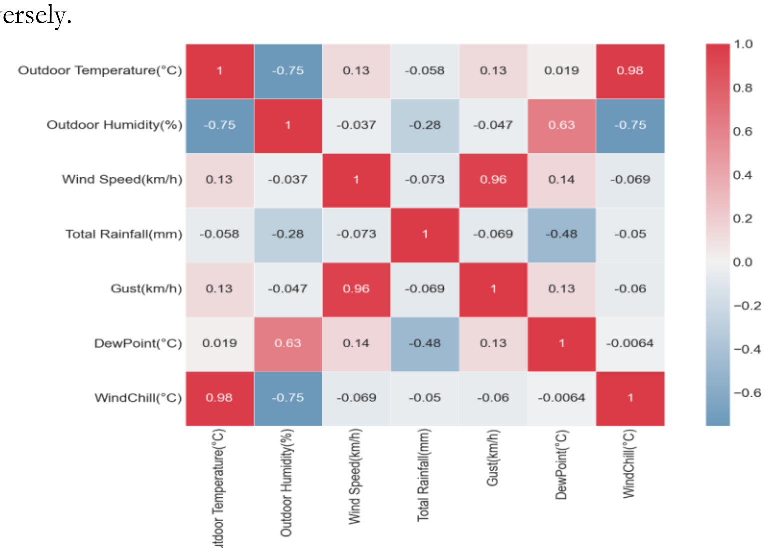

Figure 5. The correlation matrix of several weather features

2.3. LSTM

Long Short-Term Memory (LSTM) model is designed to preserve long-term dependencies and read any length sequence. The architecture of the LSTM model is illustrated in Figure 6.

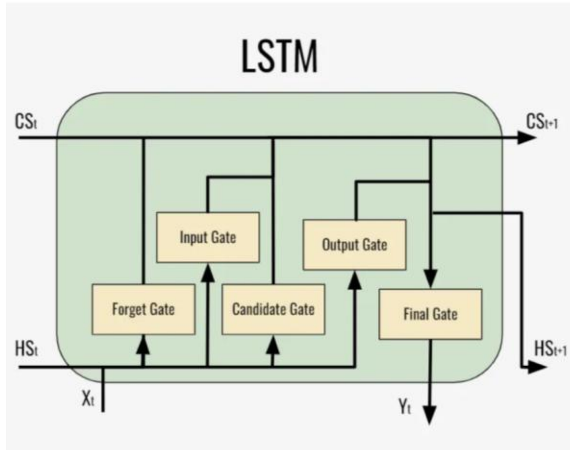

Figure 6. The architectural design of LSTM model (Hull, 2022)

As seen in Figure 6, the LSTM architecture consists of sequential blocks that repeat each other in 3 layers: forget, input and output. Here, X_t is an input; HS_t is the hidden state which is used as an input; CS_t is cell state which corresponds to the new information, and Y_t is the final state, that is the final value of the network. The goal of the forget gate is precisely to discard or erase information. The forget gate employs the sigmoid function to determine which information to retain and which to discard. The candidate gate, an activation layer, in an LSTM is essential for determining the possible new information that can be included in the cell state. This component, in conjunction with the forget gate and input gate, plays a crucial role in controlling the transmission of information within the LSTM cell. This control is essential for capturing and preserving long-term dependencies in sequences. Similar to the forget gate, the input gate determines the amount of information to be retained during different time steps of the network. The main difference lies in the fact that the input gate determines whether information entering the network should be kept, rather than determining the amount of information to be retained from memory. The input gate chooses the information to retain by multiplying significant elements with values approaching one and insignificant elements with values approaching zero. The output gate is not the last gate of the network. This gate recalls information between hidden states. The last step, before the network returns its final result, the value received from the output gate goes through the final gate.

The architecture of the LSTM model proposed for forecasting air temperature for studied area consists of; one input layer consisting of almost one month of air temperature values, a hidden layer consisting of 50 neurons with the activation function tanh and finally a dense layer consisting of one neuron with the linear activation function. The mean squared error criterion is applied as a loss function between the relevant neuron of the input and output layers. ADAM, NADAM, and RMSprop (Era et al., 2023) were used as the optimization algorithms. Determination of optimal epoch size varies according to data set and model configuration. Also to avoid overfitting or underfitting, the value should be chosen between 50 and 200 epochs (Hochreiter and Schmidhuber, 1997; Greff et al., 2016; Goodfellow et al., 2016). Therefore, for the LSTM model used within this study 50, 100 and 200 epochs are used as if those are commonly used for moderate tasks, and the batch size as 32 in the learning process of the model. A total of 262 parameters were optimized in the proposed model.

2.4. Statistical Analysis

In this study, some statistical evaluation criteria existing in the literature were used to check the performance and accuracy of the results obtained from the LSTM algorithm proposed for air temperature forecasting. Root Mean Square Error (RMSE) is one of the statistical metrics used which generally represents the standard deviation of the forecast errors.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(T_{Fore} - T_{expe}\right)^2} \tag{1}$$

where T_{expe} is the air temperature data received from PCE-FWS 20N Air Measurement Device, T_{Fore} is the air temperature forecasted via LSTM model. RMSE quantifies the average magnitude of errors between the predicted air temperature values and the actual values measured by the air measuring device used in the study. RMSE provides a comprehensive measure of forecast accuracy by giving more weight to large errors as well as individual errors (Barhmi et al., 2024). The closer the RMSE value is to zero, the closer the air temperature values estimated by the LSTM method are to the actual values obtained in the experimental device. Thus, it shows that better estimates are made.

Within this study, to define the strength of the linear relationship between forecasted value of the model and real values, coefficient of determination (R^2) is used.

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(T_{expe} - T_{Fore}\right)^2}{\sum_{i=1}^{N}\left(T_{expe} - \bar{T}_{expe}\right)^2} \tag{2}$$

where $\bar{T}_{expe}$ is the mean of air temperature data received measurement device.

3. Results and Discussion

This study utilizes the PCE-FWS 20N Weather Station, situated in the Beykoz University Undergraduate Programs Laboratory and Ateliers Campus Building in Istanbul, Turkiye, to gather empirical atmospheric data, encompassing wind direction and speed, temperature, relative humidity, and precipitation levels. Temperature forecast is conducted via machine learning method LSTM and statistical analysis method is utilized to compare the predicted and actual temperature data. Statistical analysis conducted for optimization algorithms; ADAM (Adaptive Moment Estimation), NADAM (Nesterov-accelerated Adaptive Moment Estimation), RMSprop (Root Mean Square Propagation). The epoch size is determined as 50, 100 and 200 while the batch size is taken as 32 in the learning process of the model. The results are given in Table 3.

Table 3. Analysis results.

	Adam				Nadam				RMSprop			
	RMSE		R^2		RMSE		R^2		RMSE		R^2	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
50	0.5931	0.0869	0.9483	0.0176	0.6211	0.1394	0.9417	0.0322	0.6788	0.0577	0.9333	0.0112
100	0.5792	0.0278	0.9516	0.0049	0.5904	0.0307	0.9497	0.0054	0.6283	0.0917	0.9420	0.0191
200	0.6039	0.0350	0.9474	0.0061	0.6166	0.0445	0.9450	0.0078	0.6025	0.0331	0.9476	0.0058

Table 3 reports the average (mean μ and standard deviation σ) across 10 independent random-seed initializations i.e., for each optimization algorithm for determined epoch sizes. Each optimization algorithm is executed 10 times for one epoch number. For instance, the LSTM model is conducted for air temperature forecasting with ADAM optimization algorithm with 50 numbers of epochs which is executed 10 times to define the performance of the learning process. Among the evaluated configurations, ADAM with 100 epochs yielded the lowest mean RMSE and the highest mean R^2. Therefore, this configuration was selected for the subsequent analysis and visualization. The result of the analysis is done to make an estimation for the period covered by the data set given in; Figure 7.

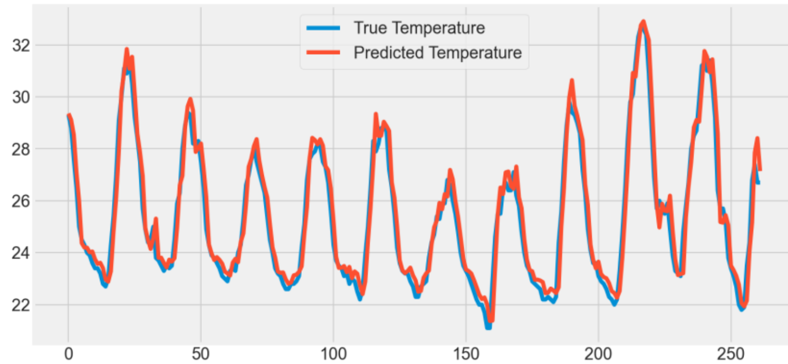

Figure 7. Temperature Forecast between July 7-September 1

The part of the result is shown in Figure 8, with columns for Date, True Temperature, and Predicted Temperature, presents the predictions from a single representative trained model instance. It is seen that the difference between predicted and actual temperature data is very slight.

	Date	True_Temperature	Predicted_Temperature
0	2023-08-21 16:00:00	29.3	29.360027
1	2023-08-21 17:00:00	28.8	29.111685
2	2023-08-21 18:00:00	27.6	28.541975
3	2023-08-21 19:00:00	26.5	27.103748
4	2023-08-21 20:00:00	25.0	25.991693
..	...	...	...
257	2023-09-01 09:00:00	24.4	24.053490
258	2023-09-01 10:00:00	26.8	25.168844
259	2023-09-01 11:00:00	27.4	27.848953
260	2023-09-01 12:00:00	26.7	28.408375
261	2023-09-01 13:00:00	26.7	27.141092

Figure 8. Actual and Expected Temperature

Table 3 reports the mean (μ) and standard deviation (σ) metrics over 10 independent trial runs to evaluate optimizer stability. The model achieved an RMSE of 0.6275 and an R^2 of 0.9434 on the test subset, closely aligning with the overall ensemble mean ($\mu_{\text{RMSE}} = 0.5792$, $\mu_{R^2} = 0.9516$). Figures 7 and 8, in contrast, present one individual run for visualization purposes. Therefore, the RMSE and R^2 values shown in the figures are not expected to be identical to the average values reported in Table 3.

It is seen that; the obtained prediction performance is consistent with previous studies reporting the successful application of machine learning techniques for air temperature forecasting. However, direct numerical comparison among studies is challenging because datasets, climatic conditions, forecasting horizons, input variables, and model architectures differ considerably. Therefore, the results of this study should be interpreted within the scope of the analyzed dataset.

4. Conclusions

Air temperature forecasting plays a critical role in energy management, agricultural planning, and public health initiatives. Additionally, it is crucial for evaluating and assessing the impacts of climate change, which serves as the primary motivation for this study. Within the scope of this study; real-world atmospheric data is collected via the PCE-FWS 20N Weather station located in Beykoz University campus building. A Long Short-Term Memory (LSTM) machine learning model is applied to forecast air temperature. Within the method, ADAM (Adaptive Moment Estimation), NADAM (Nesterov-accelerated Adaptive Moment

Estimation), RMSprop (Root Mean Square Propagation) optimization algorithms are used separately for different epoch sizes 50, 100 and 200. Each optimization algorithm is executed 10 times for each epoch number. The coefficient of determination (R^2) is used to define the strength of the linear relationship between forecasted and actual values, while the root mean square error (RMSE) is employed to evaluate the model's performance statistically. The most accurate result is gathered from the ADAM optimization algorithm with 100 numbers of epochs. Therefore, the evaluation is conducted further for only one time of execution. The study demonstrates that the LSTM machine learning model, utilizing the ADAM optimization algorithm with 100 epochs, is a reliable approach for temporal temperature forecasting.
A limitation of the current experimental design is the short timeframe of the collected dataset, spanning two months during the summer season. While this period captures peak summer temperatures critical for short-term thermal stress analysis, it does not account for multi-seasonal variations or long-term climate dynamics. For future research, this model may be explored over longer time periods across four seasons and compared with alternative forecasting methods to further validate its effectiveness.

Declaration of Ethical Standards

As the authors of this study, we declare that this study complies with all relevant ethical standards.

Credit Authorship Contribution Statement

I.B.T.D.: Data Acquisition and Curation, Writing – review & editing, Methodology, Investigation, Conceptualization, Formal analysis.
G.T.E.: Writing – review & editing, Methodology, Investigation, Software, Formal analysis

Declaration of Competing Interest

The authors declared that they have no conflict of interest.

Funding / Acknowledgements

The author/authors do not wish to acknowledge any individual or institution.

This research received no external funding.

This study does not involve human or animal participants. All procedures followed scientific and ethical principles, and all referenced studies are appropriately cited.

The authors declare that no Artificial Intelligence (AI) tools were used in the creation of this article.

Data Availability

The atmospheric dataset collected and analyzed during the current study is available from the corresponding author upon reasonable request.

References

Apaydın, M., Yumuş, M., Degirmenci, A., & Karal, Ö. (2022). Evaluation of Air Temperature with Machine Learning Regression Methods using Seoul City Meteorological Data. Pamukkale Universitesi Muhendislik Bilimleri Dergisi, 28(5), 737-747.

Bao, W., Yue, J., & Rao, Y. (2017). A Deep Learning Framework for Financial Time Series using Stacked Autoencoders and Long-Short Term Memory. Plos One, 12(7), e0180944. https://doi.org/10.1371/journal.pone.0180944.

Barhmi, K., Heynen, C., Golroodbari, S., & Van Sark, W. (2024, February). A Review of Solar Forecasting Techniques and the Role of Artificial Intelligence. In Solar (Vol. 4, No. 1, pp. 99-135). MDPI. https://doi.org/10.3390/solar4010005

Climate Change Beykoz (2024 September 17). Meteoblue. https://www.meteoblue.com/en/climate-change/beykoz_republic-of-t%c3%bcrkiye_750662?month=7

Climate Data (2024 September 17). https://tr.climate-data.org/asya/tuerkiye/istanbul/beykoz-25629/

Dadhich, S., Pathak, V., Mittal, R., & Doshi, R. (2021). Machine Learning for Weather Forecasting. Machine Learning for Sustainable Development, 10, 9783110702514-010. https://doi.org/10.1515/9783110702514-010 .

Deif, M. A., Solyman, A. A., Alsharif, M. H., Jung, S., & Hwang, E. (2021). A Hybrid Multi-Objective Optimizer-based SVM Model for Enhancing Numerical Weather Prediction: A Study for the Seoul Metropolitan Area. Sustainability, 14(1), 296. https://doi.org/10.3390/su14010296.

Dritsas, E., Trigka, M., & Mylonas, P. (2022, November). A Multi-Class Classification Approach for Weather Forecasting with Machine Learning Techniques. In 2022 17th International Workshop on Semantic and Social Media Adaptation & Personalization (SMAP) (pp. 1-5). IEEE. 10.1109/SMAP56125.2022.9942121.

Era, C. A. A., Rahman, M., & Alvi, S. T. (2023, July). Short Term Weather Forecasting Comparison Based on Machine Learning Algorithms. In 2023 Intelligent Methods, Systems, and Applications (IMSA) (pp. 369-374). IEEE. 10.1109/IMSA58542.2023.10217753.

Filipović, N., Brdar, S., Mimić, G., Marko, O., & Crnojević, V. (2022). Regional Soil Moisture Prediction System based on Long Short-Term Memory Network. Biosystems Engineering, 213, 30-38. https://doi.org/10.1016/j.biosystemseng.2021.11.019.

Fang, W., Zhuo, W., Yan, J., Song, Y., Jiang, D., & Zhou, T. (2022). Attention Meets Long Short-Term Memory: A Deep Learning Network for Traffic Flow Forecasting. Physica A: Statistical Mechanics and its Applications, 587, 126485. https://doi.org/10.1016/j.physa.2021.126485.

Findawati, Y., Astutik, I. I., Fitroni, A. S., Indrawati, I., & Yuniasih, N. (2019, December). Comparative Analysis of Naïve Bayes, K Nearest Neighbor and C. 45 Method in Weather Forecast. In Journal of Physics: Conference Series (Vol. 1402, No. 6, p. 066046). IOP Publishing. 10.1088/1742-6596/1402/6/066046.

Fischer, T., & Krauss, C. (2018). Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions. European Journal of Operational Research, 270(2), 654-669. https://doi.org/10.1016/j.ejor.2017.11.054.

Gao, P., Qiu, H., Lan, Y., Wang, W., Chen, W., Han, X., & Lu, J. (2021). Modeling for the Prediction of Soil Moisture in Litchi Orchard with Deep Long Short-Term Memory. Agriculture, 12(1), 25. https://doi.org/10.3390/agriculture12010025.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R., & Schmidhuber, J. (2016). LSTM: A Search Space Odyssey. IEEE Transactions on Neural Networks and Learning Systems, 28(10), 2222-2232. 10.1109/TNNLS.2016.2582924.

Han, H., Choi, C., Kim, J., Morrison, R. R., Jung, J., & Kim, H. S. (2021). Multiple-depth Soil Moisture Estimates using Artificial Neural Network and Long Short-Term Memory Models. Water, 13(18), 2584. https://doi.org/10.3390/w13182584.

Hochreiter S, Schmidhuber J (1997) Long Short-Term Memory. Neural Computation, 9: 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735

Holmstrom, M., Liu, D., & Vo, C. (2016). Machine Learning Applied to Weather Forecasting. Meteorol. Appl, 10(1), 1-5.

Huang, Y., Gao, Y., Gan, Y., & Ye, M. (2021). A New Financial Data Forecasting Model using Genetic Algorithm and Long Short-Term Memory Network. Neurocomputing, 425, 207-218. https://doi.org/10.1016/j.neucom.2020.04.086.

Hull, G. (2022 December 19). Building a Neural Network Zoo From Scratch: The Long Short-Term Memory Network. https://medium.com/@CallMeTwitch/building-a-neural-network-zoo-from-scratch-the-long-short-term-memory-network-1cec5cf31b7

Jaharabi, W., Hossain, M. D., Tahmid, R., Islam, M. Z., & Rayhan, T. M. (2023). Predicting Temperature of Major Cities Using Machine Learning and Deep Learning. arXiv preprint arXiv:2309.13330. https://doi.org/10.48550/arXiv.2309.13330.

Jakaria, A. H. M., Hossain, M. M., & Rahman, M. A. (2020). Smart Weather Forecasting using Machine Learning: A Case Study in Tennessee. arXiv preprint arXiv:2008.10789. https://doi.org/10.48550/arXiv.2008.10789.

Johnstone, C., & Sulungu, E. D. (2021). Application of Neural Network in Prediction of Temperature: A Review. Neural Computing and Applications, 33(18), 11487-11498. https://doi.org/10.1007/s00521-020-05582-3.

Kareem, F. Q., Abdulazeez, A. M., & Hasan, D. A. (2021). Predicting Weather Forecasting State based on Data Mining Classification Algorithms. Asian Journal of Research in Computer Science, 9(3), 13-24. 10.9734/AJRCOS/2021/v9i330222.

Katušić, D., Pripužić, K., Maradin, M., & Pripužić, M. (2022). A Comparison of Data-Driven Methods in Prediction of Weather Patterns in Central Croatia. Earth Science Informatics, 15(2), 1249-1265. https://doi.org/10.1007/s12145-022-00792-w.

Kumar, I., Tripathi, B. K., & Singh, A. (2023). Attention-based LSTM Network-assisted Time Series Forecasting Models for Petroleum Production. Engineering Applications of Artificial Intelligence, 123, 106440. https://doi.org/10.1016/j.engappai.2023.106440.

Kumari, S., & Muthulakshmi, P. (2023). A Wide Scale Survey on Weather Prediction using Machine Learning Techniques. Journal of Information & Knowledge Management, 22(05), 2250093. https://doi.org/10.1142/S0219649222500939.

Li, Q., Zhu, Y., Shangguan, W., Wang, X., Li, L., & Yu, F. (2022). An Attention-Aware LSTM Model for Soil Moisture and Soil Temperature Prediction. Geoderma, 409, 115651. https://doi.org/10.1016/j.geoderma.2021.115651.

Moorthy, R. S., & Parameshwaran, P. (2022). An Optimal K-Nearest Neighbor for Weather Prediction using Whale Optimization Algorithm. International Journal of Applied Metaheuristic Computing (IJAMC), 13(1), 1-19. 10.4018/IJAMC.290538.

Mou, L., Zhao, P., Xie, H., & Chen, Y. (2019). T-LSTM: A Long Short-Term Memory Neural Network Enhanced by Temporal Information for Traffic Flow Prediction. IEEE Access, 7, 98053-98060. 10.1109/ACCESS.2019.2929692.

Nayak, M. A., & Ghosh, S. (2013). Prediction of Extreme Rainfall Event using Weather Pattern Recognition and Support Vector Machine Classifier. Theoretical and Applied Climatology, 114, 583-603. https://doi.org/10.1007/s00704-013-0867-3.

Parlak, B. O., & Yavasoglu, H. A. (2023). Comparison of Regression Algorithms to Predict Average Air Temperature. International Journal of Engineering Research and Development, 15(1), 312-322. https://doi.org/10.29137/umagd.1232020.

Sagheer, A., & Kotb, M. (2019). Time Series Forecasting of Petroleum Production using Deep LSTM Recurrent Networks. Neurocomputing, 323, 203-213. https://doi.org/10.1016/j.neucom.2018.09.082.

Santamouris, M., Cartalis, C., Synnefa, A., & Kolokotsa, D. (2015). On the Impact of Urban Heat Island and Global Warming on the Power Demand and Electricity Consumption of Buildings -A Review. Energy and Buildings, 98, 119-124. https://doi.org/10.1016/j.enbuild.2014.09.052.

Shao, H., & Soong, B. H. (2016, November). Traffic Flow Prediction with Long Short-Term Memory Networks (LSTMs). In 2016 IEEE region 10 conference (TENCON) (pp. 2986-2989). IEEE. 10.1109/TENCON.2016.7848593.

Sharapov, R. V. (2022, May). Using Linear Regression for Weather Prediction. In 2022 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF) (pp. 1-4). IEEE. 10.1109/WECONF55058.2022.9803493.

Shrivastava, V. K., Shrivastava, A., Sharma, N., Mohanty, S. N., & Pattanaik, C. R. (2023a). Deep Learning Model for Temperature Prediction: A Case Study in New Delhi. Journal of Forecasting, 42(6), 1445-1460. https://doi.org/10.1002/for.2966.

Shrivastava, V. K., Shrivastava, A., Sharma, N., Mohanty, S. N., & Pattanaik, C. R. (2023b). Deep Learning Model for Temperature Prediction: An Empirical Study. Modeling Earth Systems and Environment, 9(2), 2067-2080. https://doi.org/10.1007/s40808-022-01609-x.

Sofi, S. S., & Oseledets, I. (2024). A Case Study of Spatiotemporal Forecasting Techniques for Weather Forecasting. GeoInformatica, 1-24. https://doi.org/10.1007/s10707-024-00530-y.

Suleman, M. A. R., & Shridevi, S. (2022). Short-Term Weather Forecasting using Spatial Feature Attention based LSTM Model. IEEE Access, 10, 82456-82468. 10.1109/ACCESS.2022.3196381.

Wang, H., Yang, J., Chen, G., Ren, C., & Zhang, J. (2023). Machine Learning Applications on Air Temperature Prediction in the Urban Canopy Layer: A Critical Review of 2011–2022. Urban Climate, 49, 101499. https://doi.org/10.1016/j.uclim.2023.101499.

Xiao, Y., & Yin, Y. (2019). Hybrid LSTM Neural Network for Short-Term Traffic Flow Prediction. Information, 10(3), 105. https://doi.org/10.3390/info10030105.