ACCSCIENCE PUBLISHING

Asian Journal of Water, Environment and Pollution

ORIGINAL RESEARCH ARTICLE

# Artificial neural network models for evapotranspiration determination using different feature combinations

**Rab Nawaz Bashir[1,2,3], Munawar Iqbal[3], Uzair Khan[3], Muhammad Ali Shahid[4], Tahir Rashid[3], Shahid Kamal[1]*, Ng Kok Why[5], and Amjad Rehman[2]**

[1]Center for Advanced Analytics, CoE for Artificial Intelligence, Faculty of Computing and Informatics, Multimedia University Cyberjaya, Selangor Malaysia

[2]Artificial Intelligence and Data Analytics (AIDA) Laboratory, College of Computer & Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia

[3]Department of Computer Science, Faculty of Information Science, COMSATS University Islamabad Vehari Campus, Vehari, Punjab, Pakistan

[4]Department of Computer Science, NAMAL University Mianwali, Mianwali, Punjab, Pakistan

[5]Center for Digital Innovations, CoE for Immersive Experience, Faculty of Computing and Informatics, Multimedia University, Cyberjaya, Selangor, Malaysia

## Abstract

Evapotranspiration (ET) plays a major role in hydrological management. ET determination is a complex task and is extensively determined using artificial neural network (ANN) models. There is a need to evaluate the potential of ANN models for ET determination using different environmental parameters in an arid region. The study evaluated the ANN models for ET determination with different weather features. A weather dataset from an arid climate in Pakistan was used to assess the performance of ANN models with different feature combinations. The ANN model performed best with four features: daily minimum temperature, daily mean temperature, daily mean relative humidity, and wind speed, compared to models trained with any other combination of features. The ANN model using these four features showed the best performance, with an $R^2$ value of 0.9899, a mean squared error of 0.0617, a mean squared error of 0.2483, and an mean absolute error of 0.1917 mm day$^{-1}$. Temperature appeared to be the most significant feature for ET determination using the ANN model at the selected location. The proposed model has potential applications in precision agriculture under arid conditions for effective drought management.

***Keywords:*** Artificial neural networks; Evapotranspiration; Water management; Temperature; Relative humidity; Wind speed

***Corresponding author:**
Shahid Kamal
(shahid.kamal@mmu.edu.my)

**Citation:** Bashir RN, Iqbal M, Khan U, *et al*. Artificial neural network models for evapotranspiration determination using different feature combinations. *Asian J Water Environ Pollut*. 2026;23(3):026010003. doi: 10.36922/AJWEP026010003

**Received:** January 2, 2026

**Revised:** February 9, 2026

**Accepted:** February 25, 2026

**Published online:** May 7, 2026



## 1. Introduction

Evapotranspiration (ET) is one of the most important components of the water cycle and of energy exchange in Earth's ecosystems.[1] It refers to water loss from the land

surface and transpiration from plants.[2] ET estimation is vital for agriculture, hydrology, and climate modelling.[3] ET modeling is also very important for precision agriculture applications in an arid climate with limited irrigation water availability.[4] Moreover, ET estimation at a given location is highly valuable for drought monitoring and management, supporting sustainable agricultural development and efficient hydrological management in irrigation-deficient regions.[5]

Forecasting ET is challenging because it depends on many weather parameters.[6] The Penman–Monteith method is the standard approach for ET determination under given meteorological conditions.[7] Because the standard method requires multiple meteorological inputs, its application is limited in regions with incomplete weather records.[8] Therefore, alternative techniques that require fewer weather inputs are necessary.[9] Due to the requirement for many weather parameters, standard approaches are difficult to apply at every location,[10] especially in regions with limited data availability. To address this situation, machine-learning (ML)-based solutions can be highly effective for ET determination using fewer weather conditions. Given the importance of machine learning and the limitations of weather conditions, there is an urgent need to explore the performance of ML models under varying weather conditions in an arid climate.[11,12]

Many ML approaches have been proposed to address the complexity of ET determinations across geographical locations. These solutions are often tailored to specific climate conditions. The intricate weather conditions and ET relationships also demand region-specific ET modeling under limited weather data. This limitation underscores the importance of examining the impact of different weather conditions on ET determination in arid regions. The study proposes an ET model for an arid region of Pakistan to support sustainable agricultural development. The selected area is an intensively cultivated agricultural region, but it suffers from irrigation water deficits. Limited irrigation water reduces agricultural productivity in the selected area. Moreover, the application of weather data for ET determination is challenging due to limited resources. Therefore, there is a need for a solution that can model ET using limited weather data.

An artificial neural network (ANN)-based model is proposed for ET predictions using multiple feature combinations to address the objectives:

(i) To evaluate multiple combinations of meteorological features, such as daily maximum temperature ($T_{max}$), daily minimum temperature ($T_{min}$), daily mean temperature ($T_{mean}$), daily mean relative humidity ($RH_{mean}$), and wind speed ($WS$), as inputs to the ANN model.
(ii) To assess the performance of the ANN model with different feature combinations.
(iii) To explore the impact of different features on the performance of the ANN model in ET prediction.

The following questions are the focus of the study:

(i) What is the performance of the ANN model when different feature combinations are used as input features?
(ii) What meteorological factor has a significant impact on ET prediction using the ANN model?

To serve these objectives, the study proposes ANN-based modeling of ET with performance evaluation of ET determination using different weather parameters. The study aims to explore the performance of the ANN model and to identify the optimal weather parameters for ET determination in selected arid regions of Pakistan.

## 2. Related work

In recent years, new ML methods for estimating ET under different weather conditions and locations have been discussed. Musanase *et al.*[1] used ET regression models on 31 years of climate data, showing that gradient boosted tree models outperformed conventional regression methods. Singh *et al.*[8] used daily meteorological features to evaluate ML models for daily reference ET prediction in India, and their results indicated that the Cubist model performed better for estimating daily ET in the selected region. According to a thorough analysis, extensive ML-oriented solutions for ET prediction have emerged in recent years. Existing ML methods focus on comparing the performance of different ML models. However, weather conditions strongly influence ET rates, and selecting appropriate features is essential for improving ET prediction accuracy. The existing literature is limited in its exploration of the impact of features on ML model performance. While algorithm performance is essential, it is equally important to assess the influence of different input features on ML model performance.

Xiang *et al.*[13] compared nine ET estimation methods in Bosnia and Herzegovina and found that the calibrated Hargreaves–Samani approach (HC) was the most accurate across various climatic zones and outperformed the standard HC approach. Lin *et al.*[14] compared the Penman–Monteith and HC methods for ET estimation in China and reported that the Penman–Monteith algorithm outperformed the HC technique for daily ET prediction. Kaissi *et al.*[15] evaluated the performance of four machine learning models using daily meteorological data from 2001 to 2020 in India. By analyzing the effects of different

climatic situations, the study explored a range of feature combinations. Wind speed and solar radiation have been identified as the most important predictors of ET, and the support vector machine model achieves the highest accuracy.

Lee *et al.*[16] evaluated thirty actual ET models in Korea and demonstrated that models using temperature and radiation data generally performed well, even with limited meteorological data. Su *et al.*[17] compared local and temporal scales in ET's cotton-producing regions between 1960 and 2019 and reported that ET was strongly influenced by climate factors, specifically maximum air temperature and wind speed. These variations have significant implications for water resources and pose challenges for cotton production under rising temperatures and drought.

Triana-Madrid *et al.*[18] proposed a monthly reference ET model for southwestern Colombia using climatic data from 1983 to 2017 and applied an ML approach to address limited climatic data availability. A study by Feng *et al.*[19] explored the relationship between soil moisture and ET. Khairan *et al.*[20] presented a reference ET prediction using an ANN combined with a Particle Swarm Optimization Grey Wolf Optimizer Algorithm (PSOGWO) hybrid algorithm. This study focused on how well the PSOGWO–ANN model identified monthly ET. The results showed that PSOGWO–ANN models performed well at predicting ET.

Tausif *et al.*[21] proposed a federated learning approach to ET determination across different locations in Pakistan. Fong *et al.*[22] estimated ET using hybrid deep-learning models trained on remote sensing data. Mustapha *et al.*[23] compared ET determination performance using ground-based data and gridded climate datasets.

Many studies have proposed ET determination using ML models. A limited number of studies focus on determining the performance of ANN models under different weather parameters. Moreover, ET determination with limited weather parameters in arid regions such as Pakistan remains scarce. There is a need to focus on ANN model performance and the impact of different weather parameters on ML-assisted ET modeling in arid regions.

## 3. Materials and methods

The study intends to explore the performance of the ANN model using different weather features as inputs. It also aims to investigate the impact of different environmental conditions on the ANN model's performance in accurately predicting ET in the arid regions of Pakistan.

### 2.1. Dataset

Weather data from 2001 to 2022 for Pakistan were used to train and evaluate the ML model. The study area was Bahawalpur, situated in southern Pakistan at a latitude of 29.3956° N and a longitude of 71.6753° E. The location of the experiment area on the world map is shown in Figure 1. The climate of the selected area is arid and hot, classified as *BWh* under the Köppen–Geiger climate classification. High temperature, low humidity, and desert storms are the major characteristics of the selected area. The selected area is very dry, with an average annual rainfall of approximately 150–200 mm.

The weather data were obtained from the National Aeronautics and Space Administration (NASA) data access source.[24] These weather data were used to determine ET using the FAO–Penman–Monteith method. Daily weather parameters were used as features, including $T_{max}$, $T_{min}$, $T_{mean}$, $RH_{mean}$, and $WS$. The $T_{mean}$ was calculated using **Equation 1** from the daily $T_{max}$ and $T_{min}$.

$$T_{mean} = \frac{T_{max} + T_{min}}{2} \tag{1}$$

where $T_{max}$ is the maximum temperature, and $T_{min}$ is the minimum temperature observed during a day after each hour. **Equation 2** to calculate the daily mean relative humidity from hourly relative humidity ($RH[t]$) observations over a 24-h period.

$$RH_{mean} = \frac{1}{24}\sum_{i=1}^{N}\int_{t_{i-1}}^{t_{i-}} RH\left(t\right)\, dt \tag{2}$$

The daily maximum WS was obtained from the daily wind speed observations recorded at one-hour intervals during a day, as expressed in **Equation 3.**

$$WS = \max\left\{WS_1, WS_2, WS_3 \ldots . WS_n\right\} \tag{3}$$

where $WS$ represents the daily maximum wind speed at time t, and $WS_i$ is the wind speed at a given time instance. The parameters $T_{max}$, $T_{min}$, $T_{mean}$, $RH_{mean}$, and $WS$ were used as weather inputs to the Penman–Monteith method for ET determination. The ET and features analysis are shown in Figure 2, while the monthly distributions of the features and ET are shown in Figure 3. The complex interrelationships among ET and all features are difficult to model using linear approaches. Therefore, there is a need to explore these hidden relationships using a data-driven,

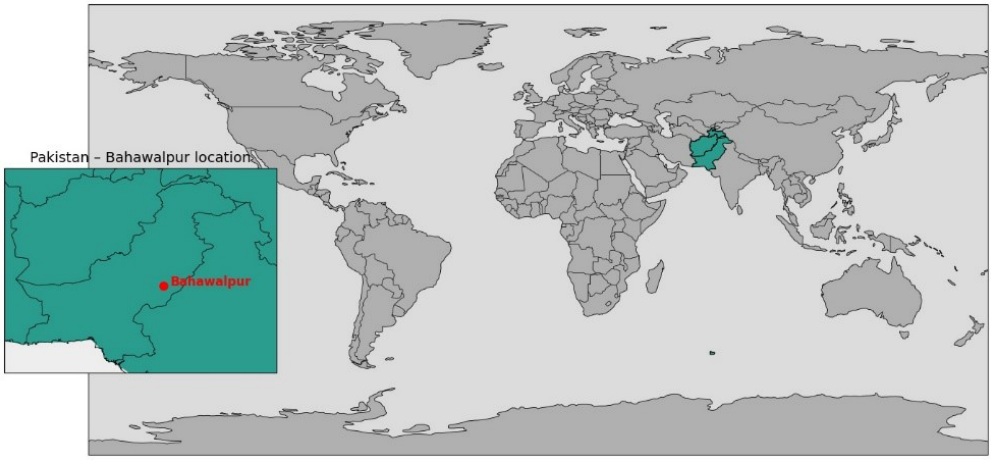


**Figure 1.** Location of the experimental area

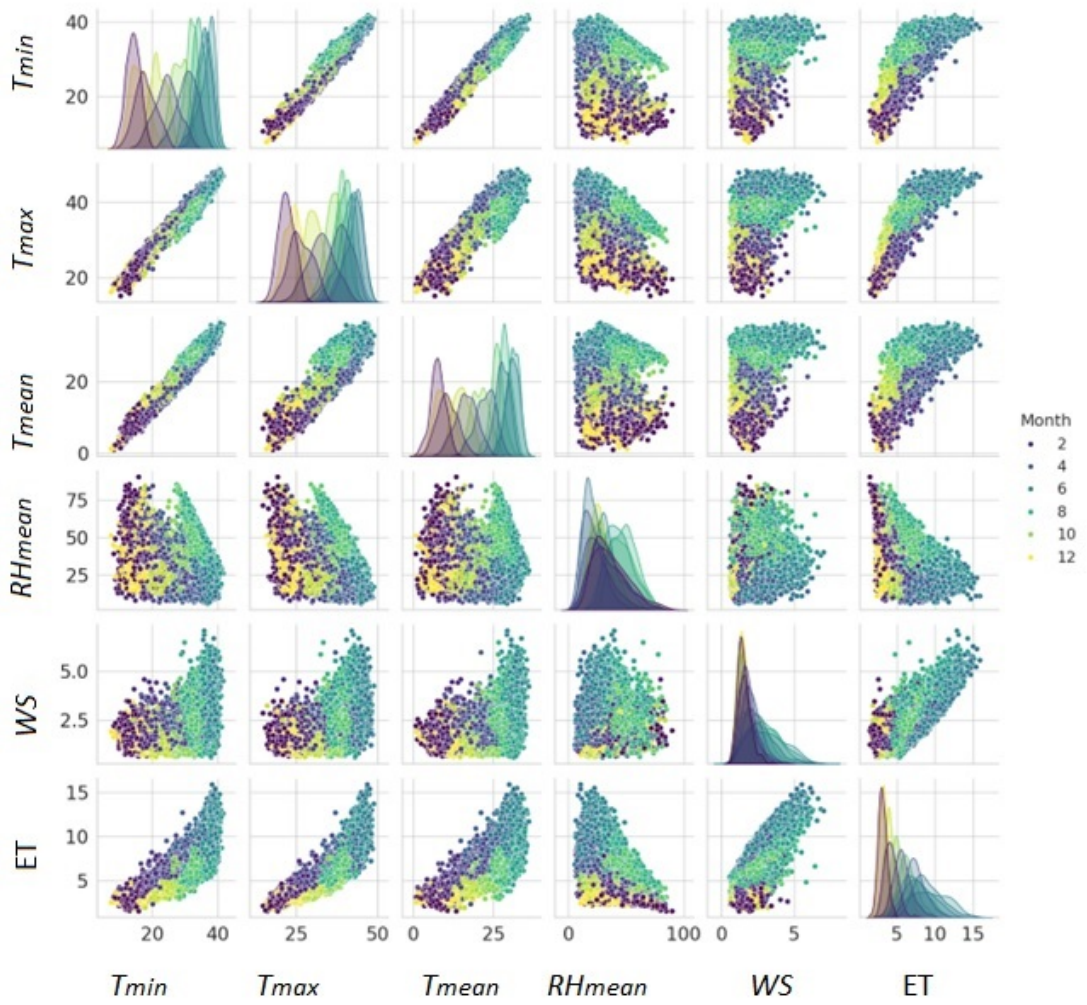


**Figure 2.** Features set and ET for each month
Abbreviations: ET: Evapotranspiration; $RH_{mean}$: Mean relative humidity; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; *WS*: Wind speed.

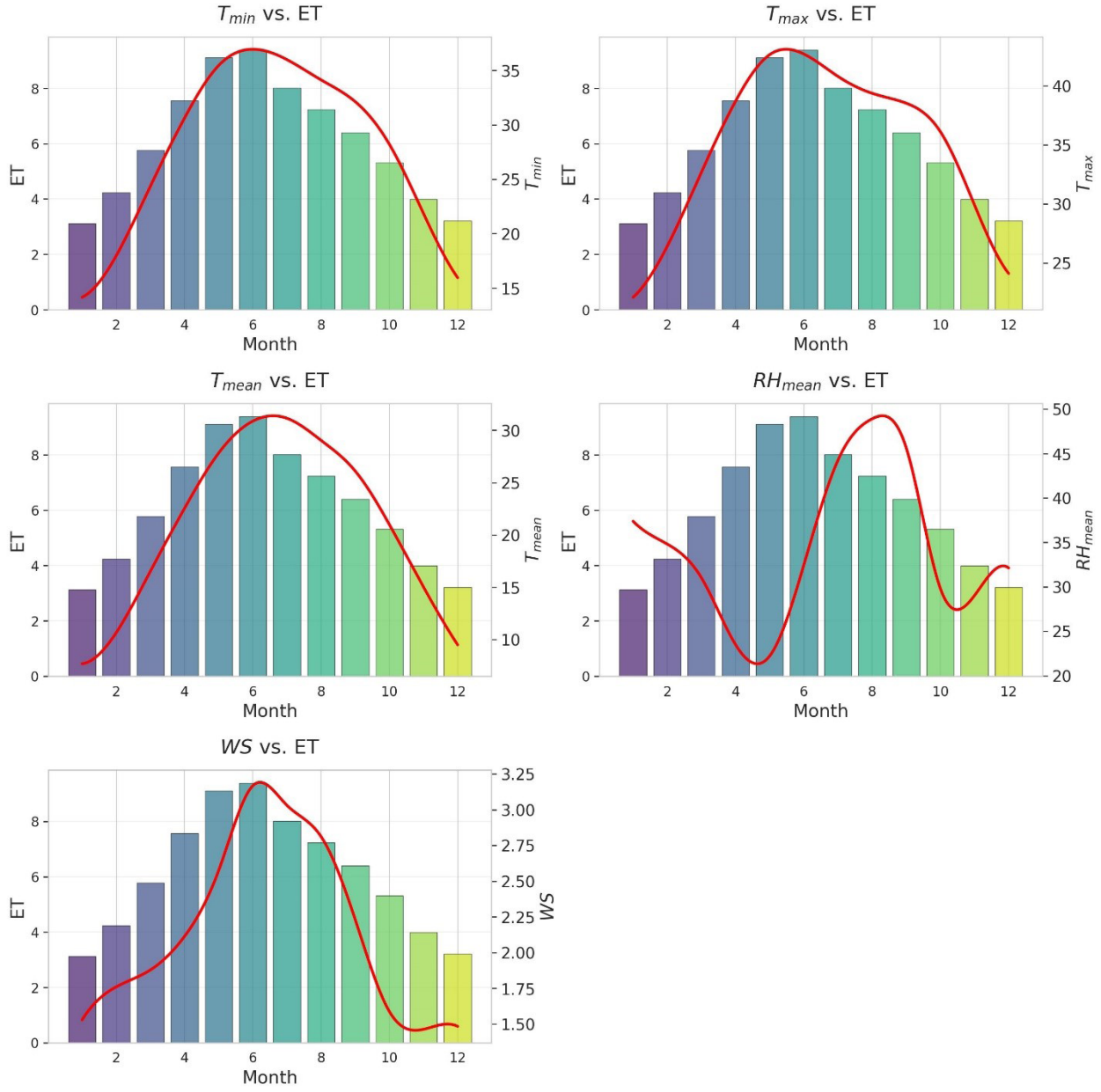


**Figure 3.** Distribution of the ET to the features for each month
Abbreviations: ET: Evapotranspiration; $RH_{mean}$: Mean relative humidity; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; *WS*: Wind speed.

ML-assisted solution.

## 2.2. Data processing

The data were processed before model training to ensure data quality. Data containing missing values were identified and removed using listwise deletion to ensure dataset consistency. Outliers were identified for each feature using the interquartile range method. Each feature was ensured to be within the natural range of its values; for example, the observed environmental conditions span a humidity range of 1–100% and a temperature range of 15–50 °C. In addition, because the input variables were meteorological in nature, they were normalized using min–max scaling to support accurate ANN model convergence and to mitigate scale-related bias among input variables. This process was consistently applied to all datasets to maintain uniformity.

### 2.3. Model training

The study aims to explore the impact of different features on the performance of the ANN models for ET prediction. For this purpose, various combinations of five features were used to train and evaluate the ANN models. The sequence diagram of the recommended approach is shown in Figure 4. Algorithm 1 defines the step-by-step procedure of the recommended process. Data preprocessing was the first step, during which rows with missing data and outliers were handled. Rows containing missing data were removed from the dataset.

**Algorithm 1. Training and evaluating artificial neural network models**

Let $F$ be the set of all features.
For $k$ = 1 to 5:
Select $k$ features from $F$ as $F_k$.
Define an ANN model $f_k(F_k, \Theta_k)$, where:

- $f_k(F_k, \Theta_k)$ represents the ANN model for the feature combination $F_k$.
- $\Theta_k$ represents the parameters of the model.

Compile the model $f_k(F_k, \Theta_k)$ with the optimizer, loss function, and metrics.
Train $f_k(F_k, \Theta_k)$ on $F_k$ using the training data.
Evaluate $f_k(F_k, \Theta_k)$ on the training set for metrics.
Evaluate $f_k(F_k, \Theta_k)$ on the testing set for metrics.
Report the evaluation metrics for $f_k(F_k, \Theta_k)$.
End For

The next step was selecting feature combinations. To evaluate the ML model's performance on unseen data, the dataset was split into 80% training and 20% test sets. A separate ML model was trained and evaluated for each weather parameter combination. The study explored the performance of the ML model defined as $ET = f(X_i)$, where $X_i$ is the weather feature vector with "$i$" being the number of features and ranging from 1 to 5, using five features. For input feature combinations with the $i$-th weather parameter, the models were represented as $f\text{-}i\text{-}j(X_1) = \sigma(X_1\Theta_1)$, where $\Theta_1$ denotes the model parameters and $j$ represented a specific combination of features. For example, for input feature combinations with one weather parameter, the models were represented as $f\text{-}1\text{-}j(X_1) = \sigma(X_1\Theta_1)$, where $\Theta_1$ were the model parameters. For input feature combinations with two features, the models were represented as $f2j(X_1) = \sigma(X_1\Theta_1)$, where $\Theta_1$ are the model parameters. Following this approach, a total of 31 ANN-based models were defined using different combinations of the five available features.

### 2.4. Machine learning model

Feed-forward ANNs were implemented to predict ET using different combinations of the five input features. The ANN-based ML models were implemented with different input layers, two hidden layers, and one output layer. The same ANN configuration was applied across all experiments and evaluated to identify the best-performing models for each feature combination. The same ANN configuration was used across all experiments to ensure architectural consistency, enabling assessment of the impact of different features on ET prediction performance. Table 1 summarizes the ANN architecture and training experiments.

**Table 1. Summary of artificial neural network configurations**

| Parameter | Value |
|---|---|
| Number of hidden layers | 2 |
| Hidden layer 1 neuron | 64 neurons |
| Hidden layer 2 neuron | 32 neurons |
| Hidden layer 1 activation function | ReLU |
| Hidden layer 2 activation function | ReLU |
| Output layer neuron | 1 neuron |
| Output layer activation function | linear activation |
| Optimizer | Adam |
| Learning rate | 0.001 |
| Loss function | MSE |
| Epochs | 50 |
| Batch size | 32 |
| Feature scaling | Min–max scaler |
| Train/test split | 80%:20% |

Abbreviations: Adam: Adaptive moment estimation; MSE: Mean squared error; ReLU: Rectified linear unit.

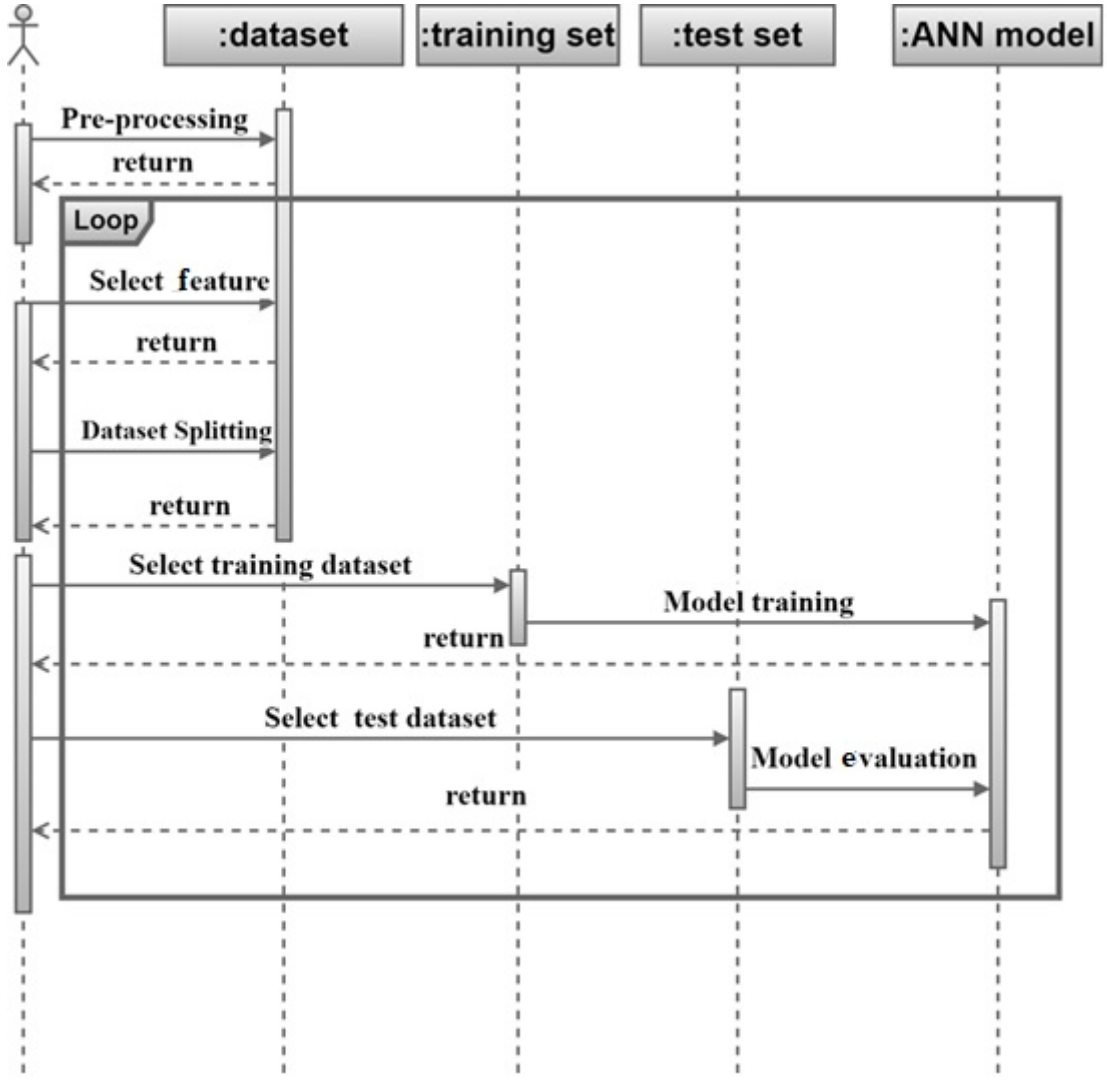


**Figure 4.** Sequence diagram of the proposed solution
Abbreviation: ANN: Artificial neural network.

The rectified linear unit activation function, shown in **Equation 4,** was used to build ANN models. Adaptive moment estimation, demonstrated in **Equation 5**, was used as a planner function. Mean squared error (MSE), expressed by **Equation 6**, served as the loss function in ANN models. In addition to MSE, the coefficient of determination ($R^2$) and mean absolute error (MAE) were also used to assess the performance of the ML models.

$$f(x) = max(0, x) \tag{4}$$

$$\theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{vt + \epsilon}} \cdot mt \tag{5}$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(target_i - prediction_i)^2 \tag{6}$$

### 2.5. Evaluation metrics

The models were evaluated using $R^2$, MSE, root mean squared error (RMSE), and MAE.

## 3. Results

In this section, the performance of the ANN models is presented. Different ANN-based machine learning models were trained using varying numbers of features. The performance of the ANN models was evaluated using 20% of the dataset as the test set. To analyze the importance

of the features, the training history of the ANN model with the top-performance features was also reported to provide insights into model performance. MSE was used as the loss function in the ANN configurations for each model. The training history of ANN models using a single feature is shown in Figure 5, and the performance of these models is summarized in Table 2. For models trained with a single feature, the ANN model using $T_{max}$ performed best, achieving an $R^2$ of 0.8058, an MSE of 1.1823 mm day$^{-1}$, an RMSE of 1.0873 mm day$^{-1}$, and an MAE of 0.7976 mm day$^{-1}$.

The training history of ANN models using two features is shown in Figure 6, and the performance of these models is summarized in Table 3. For models trained with two features, the ANN model using $T_{max}$ and $WS$ performed best, achieving an $R^2$ of 0.9496, an MSE of 0.3071 mm day$^{-1}$, an RMSE of 0.5542 mm day$^{-1}$, and an MAE of 0.4112 mm day$^{-1}$.

**Table 2. Performance comparison of the artificial neural network models with one feature**

| Feature | $R^2$ | MSE | RMSE | MAE |
|---|---|---|---|---|
| $T_{min}$ | 0.7433 | 1.5628 | 1.2501 | 0.9297 |
| $T_{max}$ | 0.8058 | 1.1823 | 1.0873 | 0.7976 |
| $T_{mean}$ | 0.6549 | 2.1009 | 1.4494 | 1.0764 |
| $RH_{mean}$ | 0.0601 | 5.7230 | 2.3923 | 1.9369 |
| $WS$ | 0.6005 | 2.4325 | 1.5597 | 1.2639 |

Abbreviations: MAE: Mean absolute error; MSE: Mean squared error; $R^2$: Coefficient of determination; $RH_{mean}$: Mean relative humidity; RMSE: Root mean squared error; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; $WS$: Wind speed.

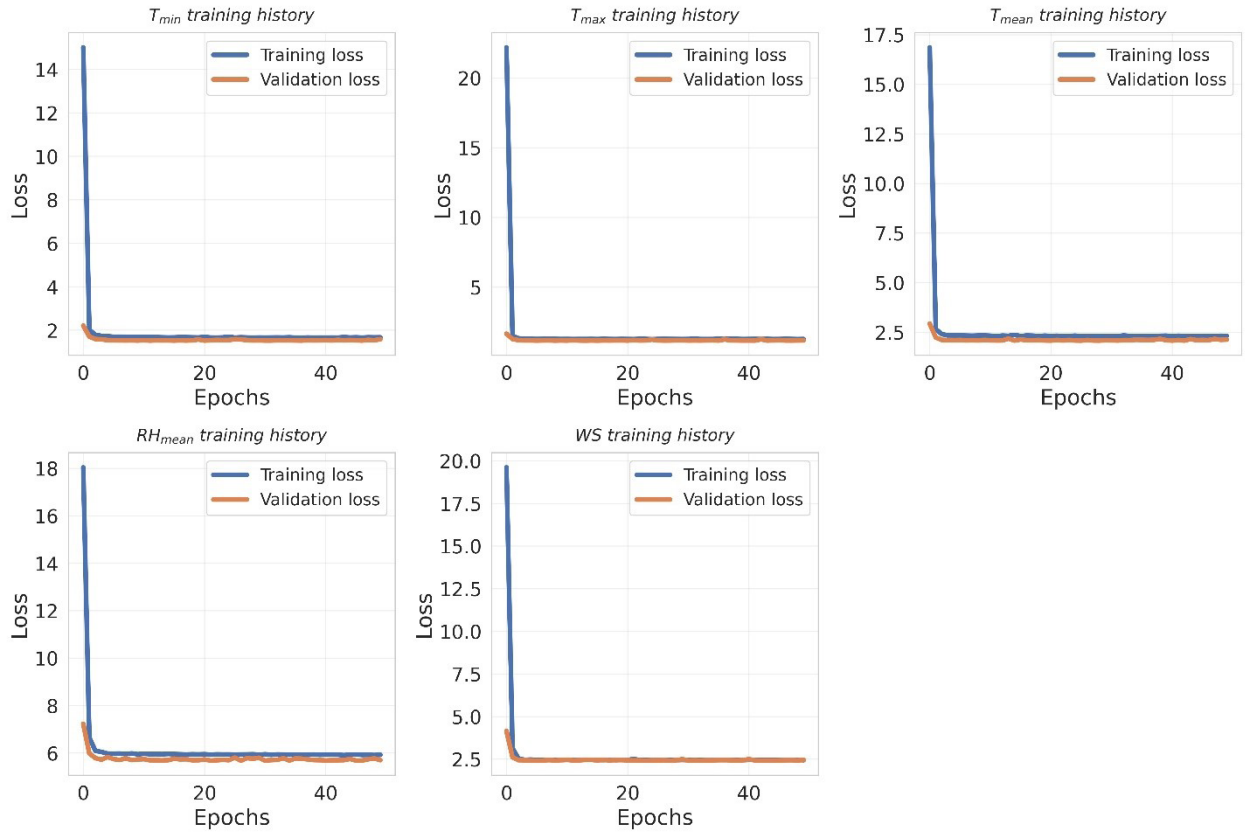


**Figure 5.** Training history of the ANN model with one feature
Abbreviations: ANN: Artificial neural network; $RH_{mean}$: Mean relative humidity; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; $WS$: Wind speed.

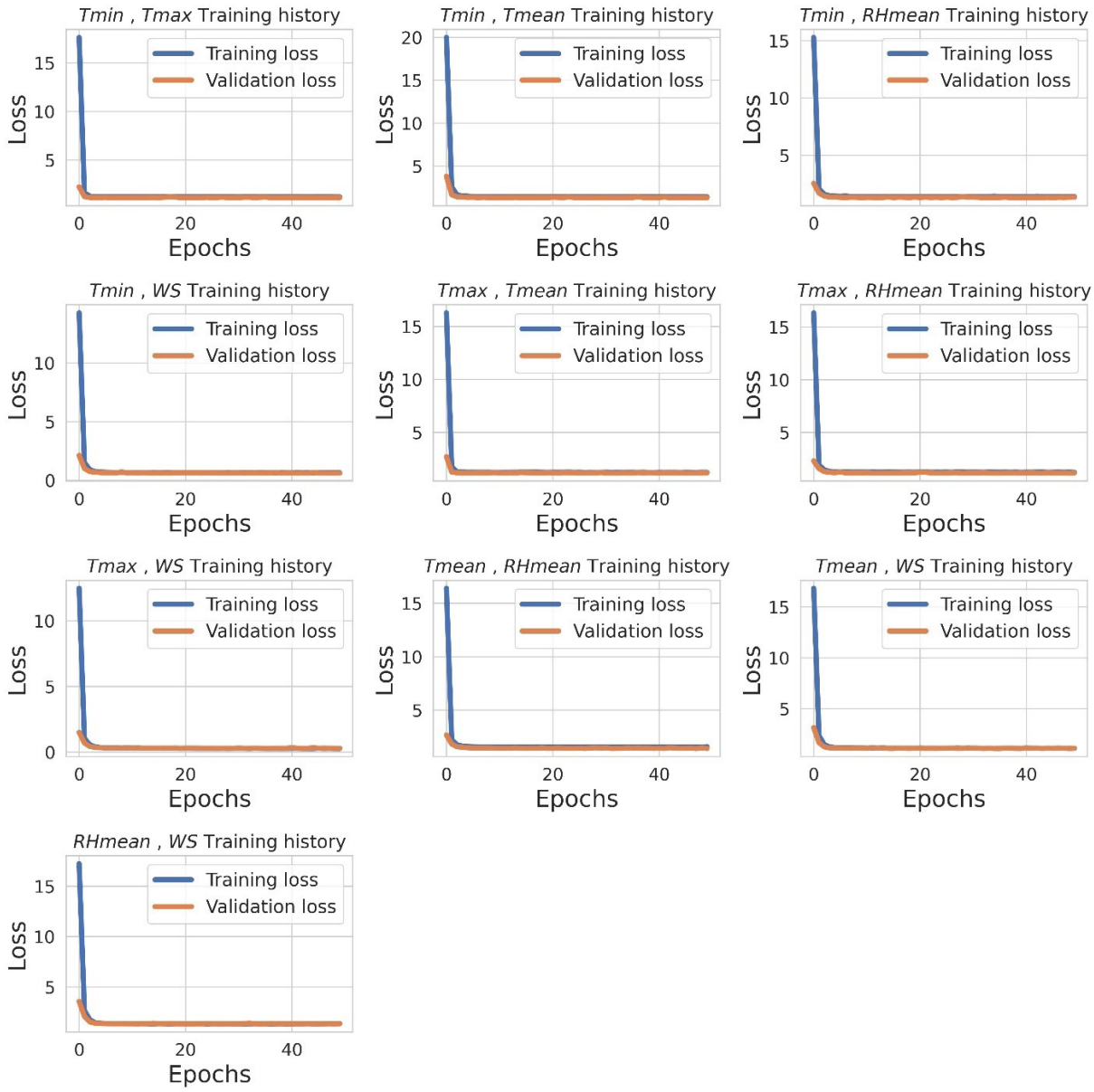


**Figure 6.** Training history of the ANN model with two features
Abbreviations: ANN: Artificial neural network; $RH_{mean}$: Mean relative humidity; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; *WS*: Wind speed.

The training history of ANN models using three features is shown in Figure 7, and the performance of these models is summarized in Table 4. For models trained with three features, the ANN model using $T_{max}$, $RH_{mean}$, and WS performed best, achieving an $R^2$ of 0.9870, an MSE of 0.0792 mm day$^{-1}$, an RMSE of 0.2815 mm day$^{-1}$, and an MAE of 0.2196 mm day$^{-1}$.

The training history of ANN models using four features is shown in Figure 8, and the performance of these models using four features is summarized in Table 5. For models trained with four features, the ANN using $T_{min}$, $T_{mean}$, $RH_{mean}$, and *WS* performed best, achieving an $R^2$ of 0.9899, an MSE of 0.0617 mm day$^{-1}$, an RMSE of 0.2483 mm day$^{-1}$, and an MAE of 0.1917 mm day$^{-1}$.

The training history of ANN models using five features is shown in Figure 9, and the performance of this model is summarized in Table 6. For the model trained with all five features, the ANN achieved an $R^2$ of 0.9879, an MSE

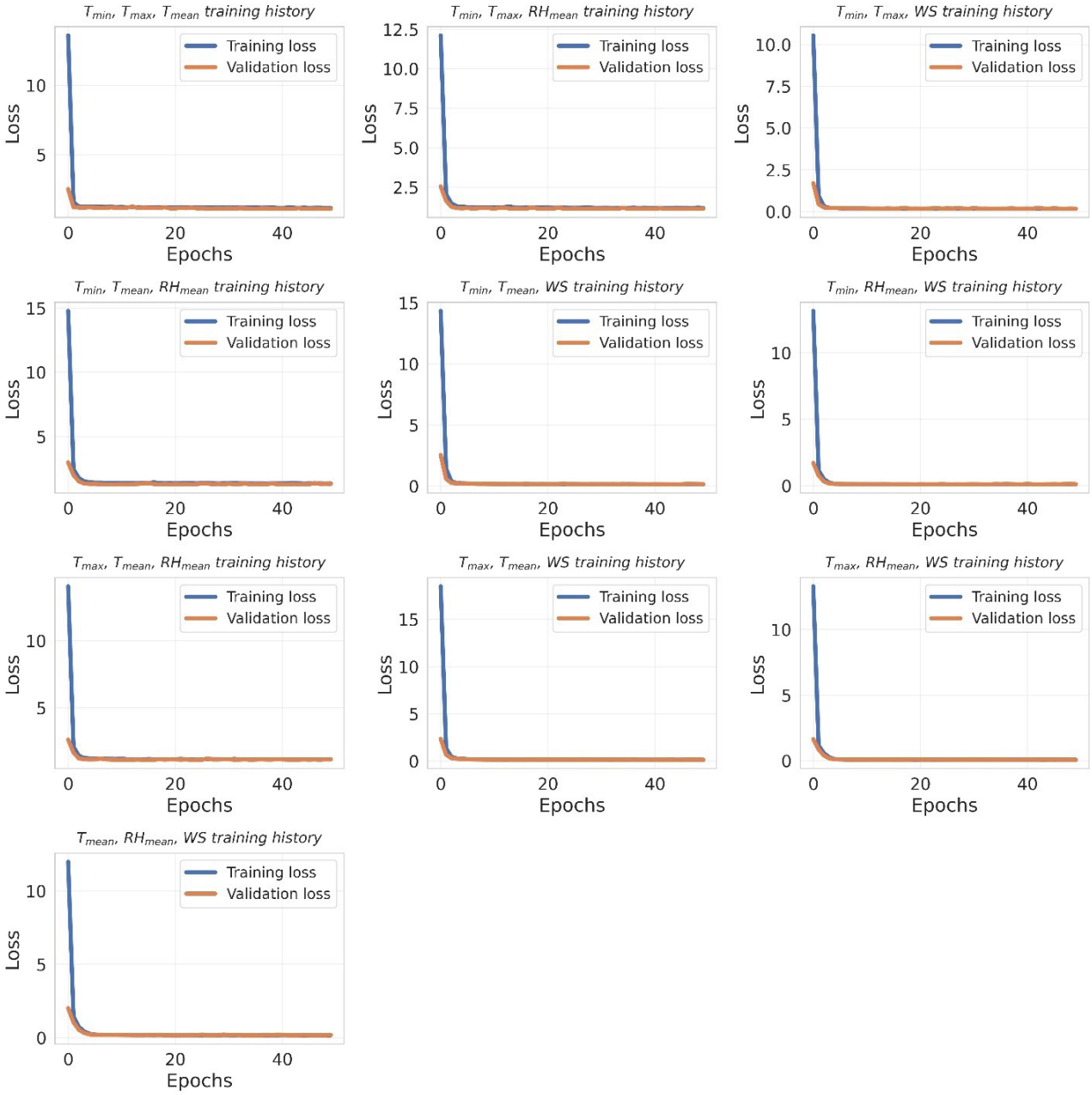


**Figure 7.** Training history of the ANN model using three features
Abbreviations: ANN: Artificial neural network; $RH_{mean}$: Mean relative humidity; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; *WS*: Wind speed.

of 0.0757, an RMSE of 0.2752, and an MAE of 0.2286 mm day$^{-1}$.

The decreasing trend in loss values over 50 epochs for each ANN model demonstrated its effective parameter optimization. The training history, as depicted by loss curves, provided insights into model convergence during training. The observed decrease in MSE across different features indicated that the iterative training process allowed the model to adapt and learn with each epoch. Over successive iterations, the model explained an increasing proportion of the variance in the target variable, as shown by the decreasing MSE scores. In addition to evaluating model performance using MSE, the top-performing ANN ML models were also compared on $R^2$, RMSE, and MAE.

**Table 3. Performance comparison of the artificial neural network model with two features**

| Feature | $R^2$ | MSE | RMSE | MAE |
|---|---|---|---|---|
| $T_{min}$ and $T_{max}$ | 0.8015 | 1.2006 | 1.0993 | 0.7993 |
| $T_{min}$ and $T_{mean}$ | 0.7824 | 1.3251 | 1.1511 | 0.8305 |
| $T_{min}$ and $RH_{mean}$ | 0.7841 | 1.3144 | 1.1465 | 0.8406 |
| $T_{min}$ and $WS$ | 0.8965 | 0.6300 | 0.7937 | 0.5879 |
| $T_{max}$ and $T_{mean}$ | 0.8113 | 1.1489 | 1.0719 | 0.7698 |
| $T_{max}$ and $RH_{mean}$ | 0.8114 | 1.1481 | 1.0715 | 0.7737 |
| $T_{max}$ and $WS$ | 0.9496 | 0.3071 | 0.5542 | 0.4112 |
| $T_{mean}$ and $RH_{mean}$ | 0.7678 | 1.4144 | 1.1891 | 0.8959 |
| $T_{mean}$ and $WS$ | 0.8185 | 1.1051 | 1.0513 | 0.7726 |
| $RH_{mean}$ and $WS$ | 0.7764 | 1.3612 | 1.1667 | 0.9473 |

Abbreviations: MAE: Mean absolute error; MSE: Mean squared error; $R^2$: Coefficient of determination; $RH_{mean}$: Mean relative humidity; RMSE: Root mean squared error; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; $WS$: Wind speed.

**Table 4. Performance comparison of the artificial neural network model with three features**

| Feature | $R^2$ | MSE | RMSE | MAE |
|---|---|---|---|---|
| $T_{min}$, $T_{max}$, and $T_{mean}$ | 0.8189 | 1.1026 | 1.0501 | 0.7587 |
| $T_{min}$, $T_{max}$, and $RH_{mean}$ | 0.8170 | 1.1088 | 1.0530 | 0.7489 |
| $T_{min}$, $T_{max}$, and $WS$ | 0.9705 | 0.1796 | 0.4238 | 0.3381 |
| $T_{min}$, $T_{mean}$, and $RH_{mean}$ | 0.7867 | 1.2987 | 1.1396 | 0.8181 |
| $T_{min}$, $T_{mean}$, and $WS$ | 0.9606 | 0.2396 | 0.4895 | 0.3781 |
| $T_{min}$, $RH_{mean}$, and $WS$ | 0.9835 | 0.1006 | 0.3171 | 0.2491 |
| $T_{max}$, $T_{mean}$, and $RH_{mean}$ | 0.8085 | 1.1660 | 1.0798 | 0.7587 |
| $T_{max}$, $T_{mean}$, and $WS$ | 0.9786 | 0.1302 | 0.3608 | 0.2793 |
| $T_{max}$, $RH_{mean}$, and $WS$ | 0.9870 | 0.0792 | 0.2815 | 0.2196 |
| $T_{mean}$, $RH_{mean}$, and $WS$ | 0.9715 | 0.1737 | 0.4168 | 0.3337 |

Abbreviations: MAE: Mean absolute error; MSE: Mean squared error; $R^2$: Coefficient of determination; $RH_{mean}$: Mean relative humidity; RMSE: Root mean squared error; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; $WS$: Wind speed.

**Table 5. Performance comparison of the artificial neural network model with four features**

| Feature | $R^2$ | MSE | RMSE | MAE |
|---|---|---|---|---|
| $T_{min}$, $T_{max}$, $T_{mean}$, $RH_{mean}$ | 0.8251 | 1.0649 | 1.0319 | 0.7362 |
| $T_{min}$, $T_{max}$, $T_{mean}$, $WS$ | 0.9785 | 0.1311 | 0.3621 | 0.2069 |
| $T_{min}$, $T_{max}$, $RH_{mean}$, $WS$ | 0.9860 | 0.0852 | 0.2919 | 0.2314 |
| $T_{min}$, $T_{mean}$, $RH_{mean}$, $WS$ | 0.9899 | 0.0617 | 0.2483 | 0.1917 |
| $T_{max}$, $T_{mean}$, $RH_{mean}$, $WS$ | 0.9883 | 0.0711 | 0.2666 | 0.2057 |

Abbreviations: MAE: Mean absolute error; MSE: Mean squared error; $R^2$: Coefficient of determination; $RH_{mean}$: Mean relative humidity; RMSE: Root mean squared error; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; $WS$: Wind speed.

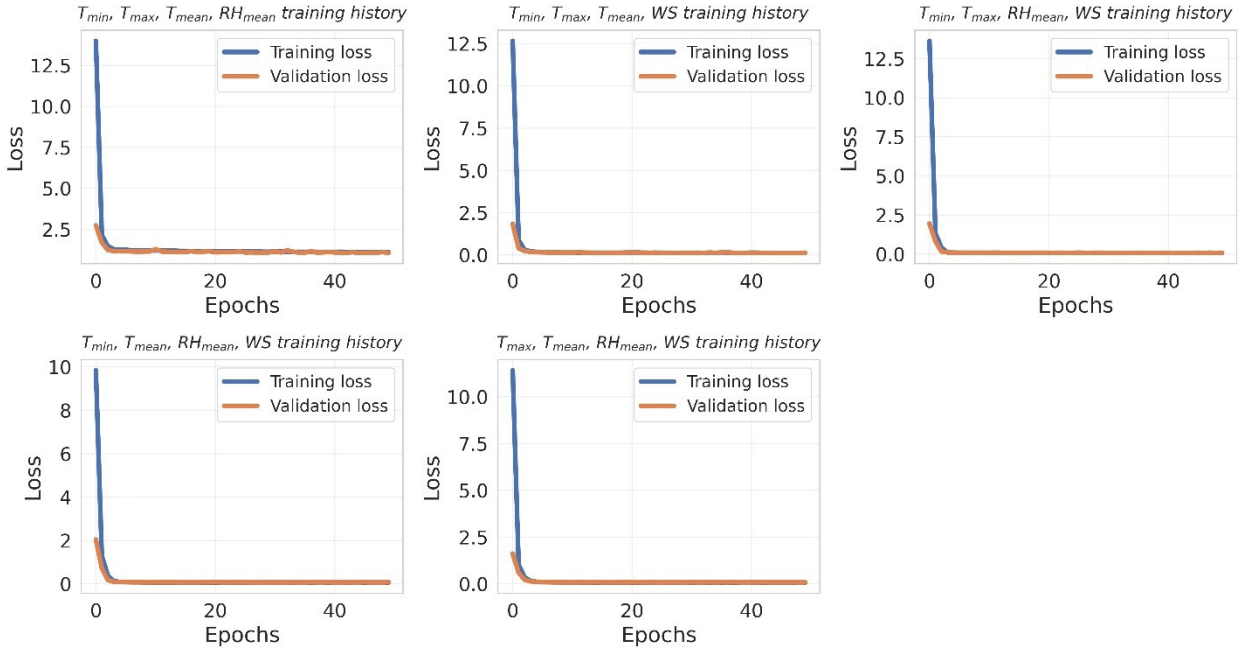


**Figure 8.** Training history of the ANN model using four features
Abbreviations: ANN: Artificial neural network; $RH_{mean}$: Mean relative humidity; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; *WS*: Wind speed.

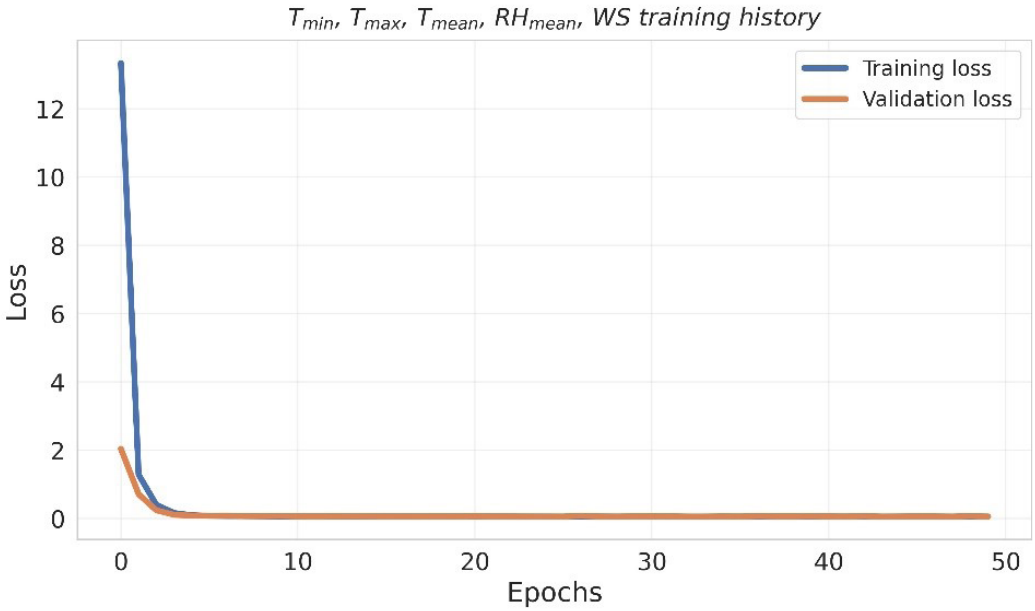


**Figure 9.** Training history of the artificial neural network model with five features

**Table 6. Performance comparison of artificial neural network machine learning models with five features**

| Features | $R^2$ | MSE | RMSE | MAE |
|---|---|---|---|---|
| $T_{min}$, $T_{max}$, $T_{mean}$, $RH_{mean}$, $WS$ | 0.9879 | 0.0757 | 0.2752 | 0.2286 |

Abbreviations: MAE: Mean absolute error; MSE: Mean squared error; $R^2$: Coefficient of determination; $RH_{mean}$: Mean relative humidity; RMSE: Root mean squared error; $T_{max}$: Maximum temperature; $T_{mean}$: Mean temperature; $T_{min}$: Minimum temperature; *WS*: Wind speed.

The performance of the top-performing ANN model for specific numbers of features is reported in Table 7. The comparison showed that the ANN model using $T_{min}$, $T_{mean}$, $RH_{mean}$, and $WS$ performed best across all models trained with different feature combinations.

The comparative analysis of $R^2$ for each top-performing feature combination is shown in Figure 10, MSE in Figure 11, RMSE in Figure 12, and MAE in Figure 13. The overall performance comparison of the ANN models with the best-performing features is summarized in Table 7.

From the ANN models with a single weather parameter ($X_1$), the model trained with $T_{max}$ exhibited the lowest MSE (1.18), RMSE (1.09), and MAE (0.80 mm day$^{-1}$), along with a relatively high $R^2$ value of 0.81. For ANN models with two features ($X_2$), the best-performing model was trained using $T_{max}$ and $WS$, achieving a substantially lower MSE of 0.31, RMSE of 0.55, MAE of 0.41 mm day$^{-1}$, and an $R^2$ value of 0.95. For three features ($X_3$), the model using $T_{max}$, $RH_{mean}$, and $WS$ outperformed other combinations, with an MSE of 0.08, an RMSE of 0.28, an MAE of 0.22 mm day$^{-1}$, and an $R^2$ of 0.99. For four features ($X_4$), the ANN model trained with $T_{min}$, $T_{mean}$, $RH_{mean}$, and $WS$ achieved the best overall performance, with an MSE of 0.06, an RMSE

**Table 7. Performance comparison of the artificial neural network model with the best-performing features**

| Features | Symbol | $R^2$ | MSE | RMSE | MAE |
|---|---|---|---|---|---|
| $T_{max}$ | $X_1$ | 0.8058 | 1.1823 | 1.0873 | 0.7976 |
| $T_{max}$, $WS$ | $X_2$ | 0.9496 | 0.3071 | 0.5542 | 0.4112 |
| $T_{max}$, $RH_{mean}$, $WS$ | $X_3$ | 0.9870 | 0.0792 | 0.2815 | 0.2196 |
| $T_{min}$, $T_{mean}$, $RH_{mean}$, $WS$ | $X_4$ | 0.9899 | 0.0617 | 0.2483 | 0.1917 |
| $T_{min}$, $T_{max}$, $T_{mean}$, $RH_{mean}$, $WS$ | $X_5$ | 0.9879 | 0.0757 | 0.2752 | 0.2286 |

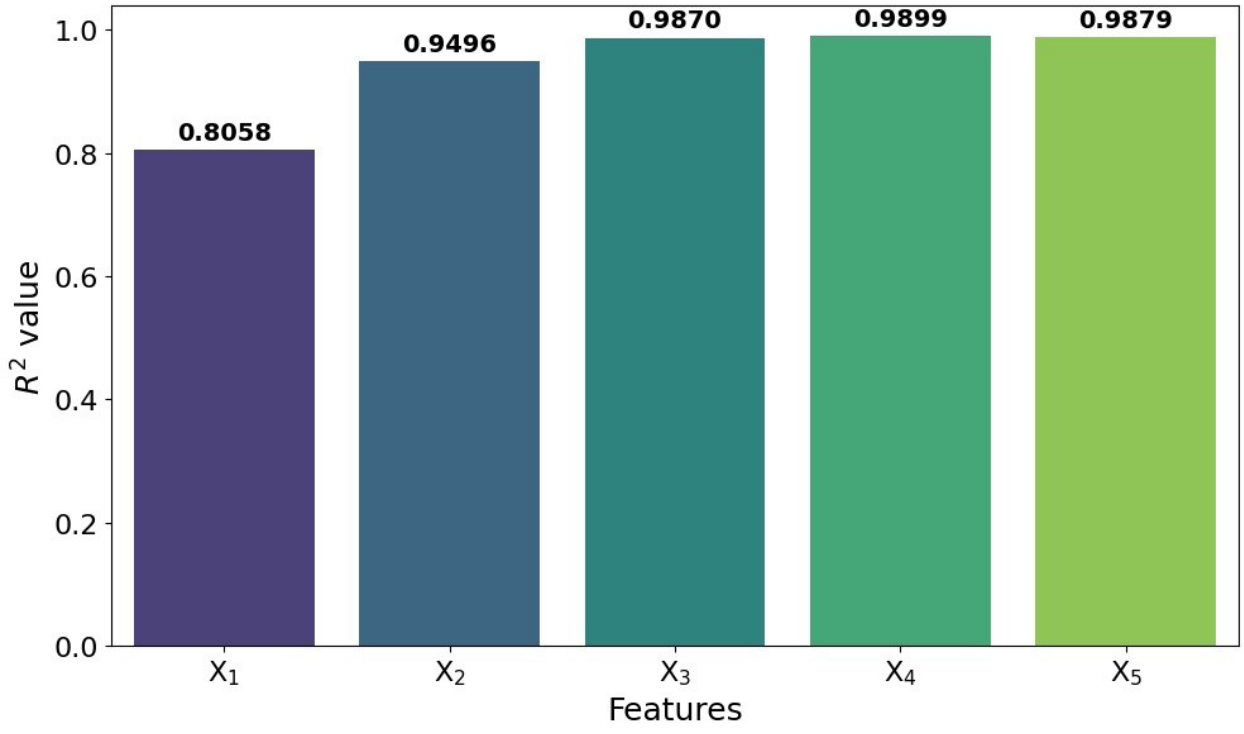


**Figure 10.** Comparison of $R^2$ of top-performing ANN models
Abbreviations: ANN: Artificial neural network; $R^2$: Coefficient of determination.

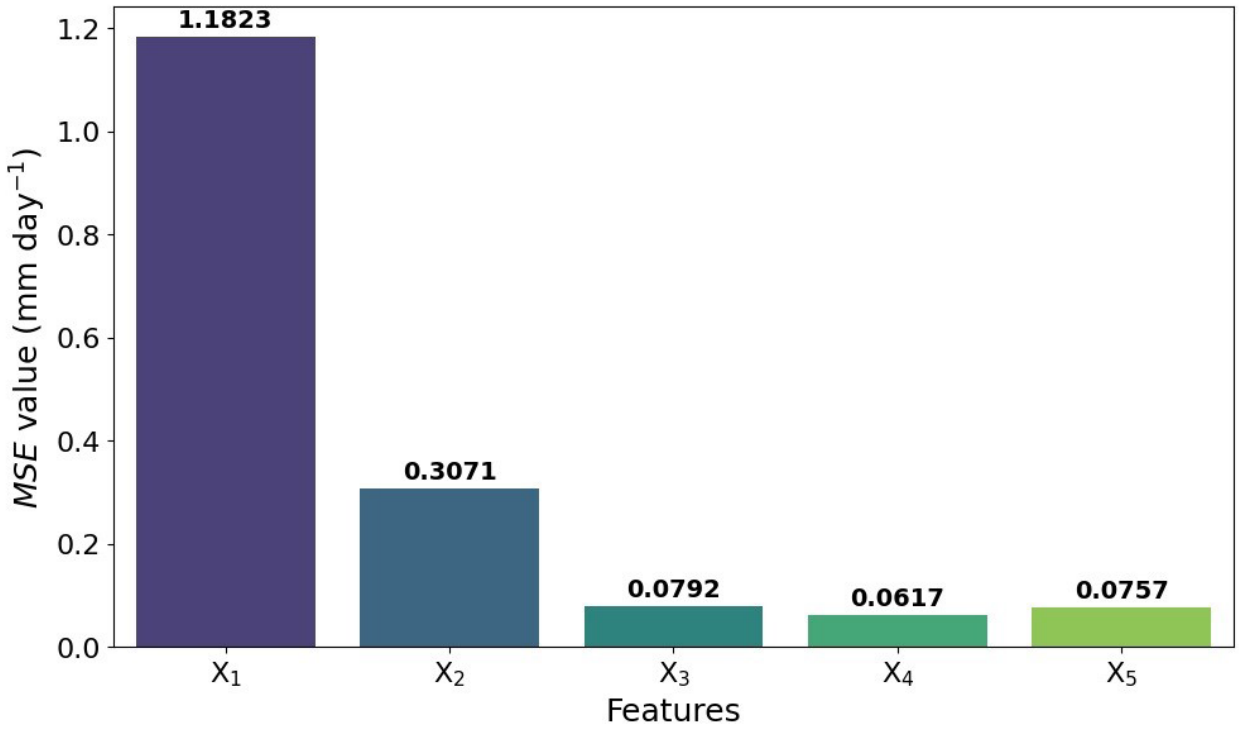


**Figure 11.** Comparison of the MSE of top-performing ANN models
Abbreviations: ANN: Artificial neural network; MSE: Mean squared error.

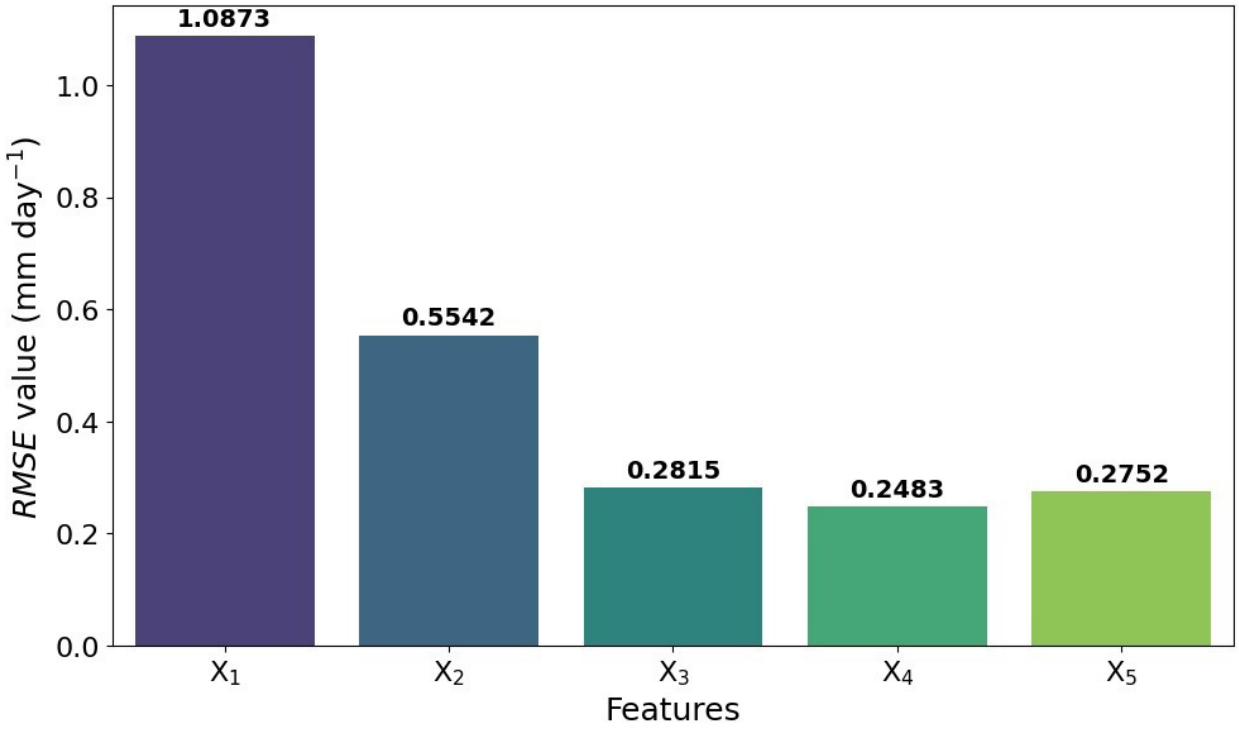


**Figure 12.** Comparison of RMSE of top-performing ANN models
Abbreviations: ANN: Artificial neural network; RMSE: Root mean squared error.

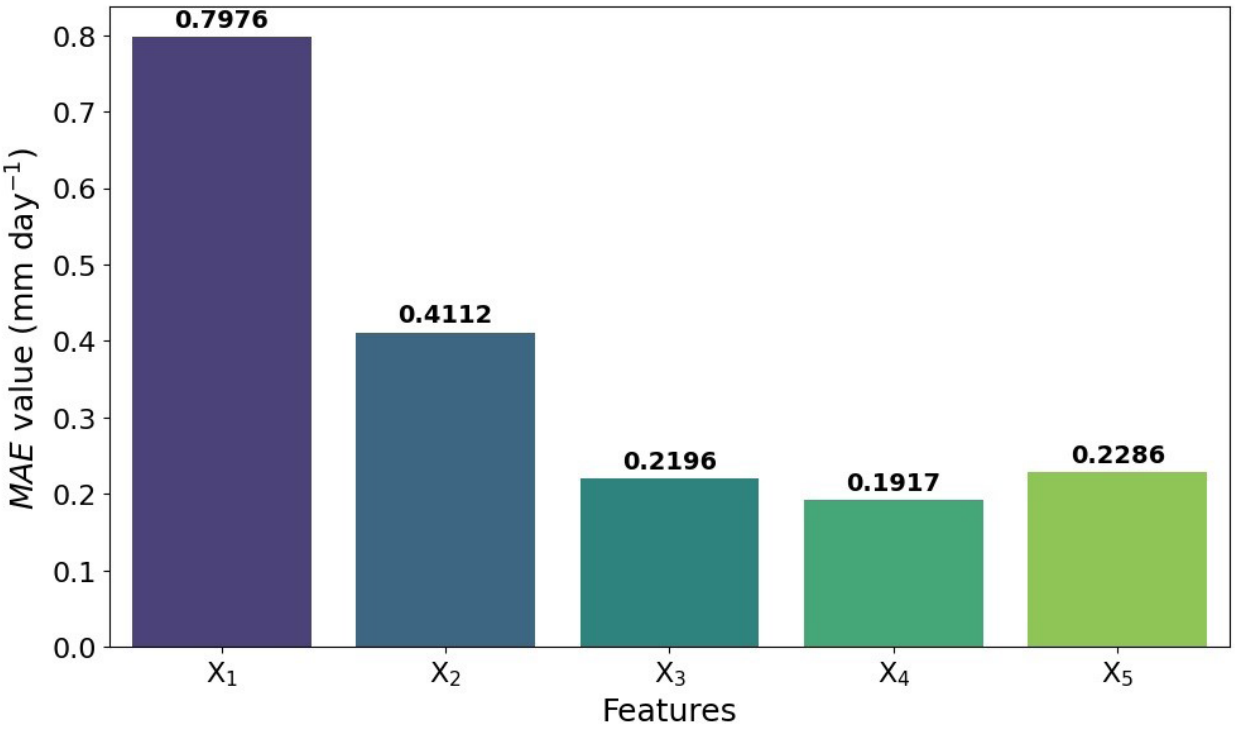


**Figure 13.** Comparison of the MAE of top-performing ANN models
Abbreviations: ANN: Artificial neural network; MAE: Mean absolute error.

of 0.25, an MAE of 0.19 mm day$^{-1}$, and an $R^2$ of 0.99. The ANN model using all five features ($X_5$) also demonstrated strong performance, with an MSE of 0.08, an RMSE of 0.28, an MAE of 0.23 mm day$^{-1}$, and an $R^2$ of 0.99. Overall, temperature emerged as the most significant feature for ET predictions in the selected area.

## 4. Discussion

The results obtained from implementing ANN models with different feature combinations provided valuable insights into the role of different features in ML-based ET estimation. The ANN trained with five features ($X_5$) achieved high ET prediction accuracy, with an $R^2$ of 0.9879, a low MSE of 0.0757, and an MAE value of 0.2286 mm day$^{-1}$. Among the four-feature combinations, the model using $T_{min}$, $T_{mean}$, $RH_{mean}$, and *WS* ($X_4$) achieved the best performance, with an $R^2$ value of 0.9899, an MSE of 0.0617, an RMSE of 0.2483, and an MAE of 0.1917 mm day$^{-1}$. This combination proved to be the most accurate ANN model for ET prediction in the selected region. For three-feature combinations ($X_3$), the model incorporating $T_{max}$, $RH_{mean}$, and *WS* performed best, achieving an $R^2$ value of 0.9870, an MSE of 0.0792, an RMSE of 0.2815, and an MAE of 0.22 mm day$^{-1}$.

For two features ($X_2$), the $T_{max}$ and *WS*-based model showed the best performance, with an $R^2$ of 0.9496, an MSE of 0.3071, an RMSE of 0.5542, and an MAE of 0.41 mm day$^{-1}$. For a single input feature (X1), the $T_{max}$-based model performed best, with an $R^2$ of 0.8058, an MSE of 1.1823, an RMSE of 1.0873, and an MAE of 0.80 mm day$^{-1}$. Across all feature combinations, temperature consistently emerged as the most influential meteorological parameter for ET prediction. The most accurate ET prediction models consistently included temperature in one or more forms ($T_{max}$, $T_{min}$, or $T_{mean}$). This predominance of temperature in ET forecasting is consistent with previous studies.[25-29] In an arid region, temperature is a dominant driver of ET compared to other weather parameters; temperature proved to be the most influential factor in ET determination. *WS*, particularly when combined with high temperature, also exerted a subtle impact on ET in arid regions. Higher wind speed, coupled with high temperature, results in a high ET rate in arid regions. Therefore, the combined use of temperature and wind speed information results in more accurate ET predictions. The inclusion of humidity further enhanced model accuracy in ET prediction, given its inverse relationship with ET. These findings correlated well with the arid and desert climate characteristics of the selected region.

Several limitations of this study should be

acknowledged. The performance evaluation of the ANN model with different weather parameters was limited to the arid climate of Pakistan. Moreover, the study examined only five meteorological variables and focused exclusively on ANN modeling. The study analyzed the ANN model's capabilities in relation to the impact of five features. There is a need to explore the performance of other ML models with more weather parameters under diverse climate conditions and in other arid regions of the world. Future work should also focus on the effects of radiation, rainfall, and atmospheric pressure on ET predictions using the ANN model. Apart from these, exploring the importance of features using other ML models and ensemble methods is also recommended for future work to enhance ET modeling accuracy.

## 5. Conclusion

Artificial neural network models demonstrated excellent performance in ET predictions, achieving an $R^2$ value of 0.99, an MSE of 0.08, and a MAE of 0.23 mm day$^{-1}$ when five features—maximum temperature, minimum temperature, mean temperature, mean relative humidity, and wind speed—were used as inputs. Among all feature combinations, the four-feature model ($X_4$), incorporating $T_{min}$, $T_{mean}$, $RH_{mean}$, and $WS$, proved to be the most accurate ANN model for ET prediction in the selected area, with an $R^2$ value of 0.9899, an MSE of 0.0617, an RMSE of 0.2483, and an MAE of 0.1917 mm day$^{-1}$. The ANN model trained with five features ($X_5$) also showed strong predictive capability, yielding an $R^2$ value of 0.9879, a low MSE of 0.0757, and an MAE of 0.2286 mm day$^{-1}$. Additionally, the ANN-based ML model performed accurately using only three features. In this case, the combination of $T_{max}$, $RH_{mean}$, and $WS$ outperformed the other three-feature combinations, achieving an $R^2$ of 0.9870, an MSE of 0.0792, an RMSE of 0.2815, and an MAE of 0.22 mm day$^{-1}$. Overall, the ANN-based ML model proved highly effective at capturing the intricate relationships underlying ET prediction, even when only a limited set of weather features was used. Temperature consistently emerged as the most influential meteorological parameter for ET prediction using an ANN model at the selected location. The proposed approach has significant implications for irrigation water management and sustainable agricultural development in water-limited regions. A major limitation of this study is the use of data from a single location. Future work should focus on evaluating the performance of ANN models and other ML approaches using additional weather parameters in some arid and more diverse climates to enhance the generalizability of the findings.

## Acknowledgments

The authors acknowledge Multimedia University, Cyberjaya, Malaysia, for providing research support and facilities. The authors are thankful to the Artificial Intelligence & Data Analytics Lab (AIDA), College of Computer and Information Sciences (CCIS), Prince Sultan University, Riyadh 11586, Saudi Arabia.

## Funding

This research is supported by Multimedia University (MMU) through its Article Page Charge (APC) Sponsorship Scheme.

## Conflict of interest

The authors declare no conflict of interest.

## Author contributions

*Conceptualization:* All authors
*Formal analysis:* All authors
*Investigation:* All authors
*Methodology:* All authors
*Writing–original draft:* All authors
*Writing–review & editing:* All authors

## Availability of data

The dataset can be requested at rabnawaz@cuivehari.edu.pk

## References

1. Musanase C, Vodacek A, Hanyurwimfura D, Uwitonze A, Kabandana I. Data-Driven Analysis and Machine Learning-Based Crop and Fertilizer Recommendation System for Revolutionizing Farming Practices. *Agriculture*. 2023;13(11):2141.

   doi: 10.3390/agriculture13112141

2. Junxu C, Jihui Z, Jiabin P, *et al*. Alp-valley and Elevation Effects on the Reference Evapotranspiration and the Dominant Climate Controls in Red River Basin, China: Insights from Geographical Differentiation. *J Hydrol (Amst)*. 2023;625:129397.

   doi: 10.1016/j.jhydrol.2023.129397

3. Zhang Y, He T, Liang S, Zhao Z. A framework for estimating actual evapotranspiration through spatial heterogeneity-based machine learning approaches. *Agric Water Manag*. 2023;289:108499.

   doi: 10.1016/j.agwat.2023.108499

4. Bashir RN, Khan FA, Khan AA, *et al*. Intelligent optimization

of Reference Evapotranspiration (ETo) for precision irrigation. *J Comput Sci*. 2023;69:102025.

doi: 10.1016/j.jocs.2023.102025

5. Hu Z, Bashir RN, Rehman AU, Iqbal SI, Shahid MMA, Xu T. Machine Learning Based Prediction of Reference Evapotranspiration (ET0) Using IoT. *IEEE Access*. 2022;10:70526-70540.

doi: 10.1109/ACCESS.2022.3187528

6. Khan AA, Nauman MA, Bashir RN, *et al*. Context Aware Evapotranspiration (ETs) for Saline Soils Reclamation. *IEEE Access*. 2022;10:110050-110063.

doi: 10.1109/ACCESS.2022.3206009

7. de Freitas EM, da Silva GH, Guimarães GFC, *et al*. Evapotranspiration and crop coefficient of Physalis peruviana cultivated with recycled paper as mulch. *Sci Hortic*. 2023;320:112212.

doi: 10.1016/j.scienta.2023.112212

8. Singh AK, Singh JB, Das B, Singh R, Ghosh A, Kantwa SR. Evaluation of machine learning models for prediction of daily reference evapotranspiration in semi-arid India. *Range Manag Agrofor*. 2023;44(1):118-125.

doi: 10.59515/rma.2023.v44.i1.14

9. Patle GT, Mandal BP, Kumar M, Jhajharia D. Modelling of Reference Evapotranspiration using Neural Network and Regression Approaches for Semi-humid Region of Sikkim. *J Agric Eng*. 2023;60(2):205-217.

doi: 10.52151/jae2023602.1808

10. Liang Y, Feng D, Sun Z, Zhu Y. Evaluation of Empirical Equations and Machine Learning Models for Daily Reference Evapotranspiration Prediction Using Public Weather Forecasts. *Water (Basel)*. 2023;15(22):3954.

doi: 10.3390/w15223954

11. Togneri R, dos Santos DF, Camponogara G, *et al*. Soil moisture forecast for smart irrigation: The primetime for machine learning. *Expert Syst Appl*. 2022;207:117653.

doi: 10.1016/J.ESWA.2022.117653

12. Bashir RN, Bajwa IS, Shahid MMA. Internet of Things and Machine-Learning-Based Leaching Requirements Estimation for Saline Soils. *IEEE Internet Things J*. 2020;7(5):4464-4472.

doi: 10.1109/JIOT.2019.2954738

13. Srdic S, Srđevic Z, Stricevic R, *et al*. Assessment of Empirical Methods for Estimating Reference Evapotranspiration in Different Climatic Zones of Bosnia and Herzegovina. *Water (Basel)*. 2023;15(17):3065.

doi: 10.3390/w15173065

14. Xiang J, Hayat M, Qiu GY, *et al*. Assessing the variations of evapotranspiration and its environmental controls over a subalpine wetland valley in China. *J Hydrol (Amst)*. 2023;617:129058.

doi: 10.1016/J.JHYDROL.2022.129058

15. Heramb P, Rao KVR, Subeesh A, *et al*. Reference evapotranspiration estimation using machine learning approaches for arid and semi-arid regions of India. *Clim Res*. 2023;91:97-120.

doi: 10.3354/cr01723

16. Lee E, Sadeghi SMM, Deljouei A, Cohen MJ. A multi-decadal national scale assessment of reference evapotranspiration methods in continental and temperate climate zones of South Korea. *J Hydrol (Amst)*. 2023;625:130021.

doi: 10.1016/j.jhydrol.2023.130021

17. Su Y, Wang J, Li J, *et al*. Spatiotemporal changes and driving factors of reference evapotranspiration and crop evapotranspiration for cotton production in China from 1960 to 2019. *Front Environ Sci*. 2023;11:1251789.

doi: 10.3389/fenvs.2023.1251789

18. Triana-Madrid JC, Ocampo-Marulanda C, Carvajal-Escobar Y, Torres-López WA, Triana J, Canchala T. Estimation of Monthly Reference Evapotranspiration with Scarce Information Using Machine Learning in Southwestern Colombia. *Meteorologica*. 2023;48(2):e024.

doi: 10.24215/1850-468Xe024

19. Feng H, Wu Z, Dong J, Zhou J, Brocca L, He H. Transpiration – Soil evaporation partitioning determines inter-model differences in soil moisture and evapotranspiration coupling. *Remote Sens Environ*. 2023;298:113841.

doi: 10.1016/j.rse.2023.113841

20. Khairan HE, Zubaidi SL, Raza SF, Hameed M, Al-Ansari N, Ridha HM. Examination of Single- and Hybrid-Based Metaheuristic Algorithms in ANN Reference Evapotranspiration Estimating. *Sustainability*. 2023;15(19):14222.

doi: 10.3390/su151914222

21. Tausif M, Iqbal MW, Bashir RN, AlGhofaily B, Elyassih A, Khan AR. Federated learning based reference evapotranspiration estimation for distributed crop fields. *PLoS One*. 2025;20(2):e0314921.

doi: 10.1371/journal.pone.0314921

22. Fong TY, Huang YF, Chin RJ, Koo CH. Enhanced estimation of reference evapotranspiration using hybrid deep learning models and remote sensing variables. *Agric Water Manag*. 2025;315:109534.

doi: 10.1016/j.agwat.2025.109534

23. Mustapha M, Zineddine M, Gmira M, El Mouhtadi M, Alaoui AEH. Machine Learning Approaches for Predicting Reference Evapotranspiration: A Comparative Study Using Ground and Gridded Climate Data in Fes Region. *World*

*Water Policy*. 2025;11(2):652-669.

doi: 10.1002/wwp2.12255

24. POWER Data Access Viewer. Available from: https://power.larc.nasa.gov/data-access-viewer/ [Last accessed on February 10, 2026].

25. Mai M, Wang T, Han Q, Jing W, Bai Q. Comparison of environmental controls on daily actual evapotranspiration dynamics among different terrestrial ecosystems in China. *Sci Total Environ*. 2023;871:162124.

doi: 10.1016/J.SCITOTENV.2023.162124

26. Wu G, Lu X, Zhao W, *et al*. The increasing contribution of greening to the terrestrial evapotranspiration in China. *Ecol Modell*. 2023;477:110273.

doi: 10.1016/J.ECOLMODEL.2023.110273

27. Ghafouri-Azar M, Lee SI. Meteorological Influences on Reference Evapotranspiration in Different Geographical Regions. *Water*. 2023;15(3):454.

doi: 10.3390/W15030454

28. Xiao C, Cai J, Zhang B, Chang H, Wei Z. Evaluation and verification of two evapotranspiration models based on precision screening and partitioning of field temperature data. *Agric Water Manag*. 2023;278:108166.

doi: 10.1016/J.AGWAT.2023.108166

29. Troncoso-García AR, Brito IS, Troncoso A, Martínez-Álvarez F. Explainable hybrid deep learning and Coronavirus Optimization Algorithm for improving evapotranspiration forecasting. *Comput Electron Agric*. 2023;215:108387.

doi: 10.1016/j.compag.2023.108387