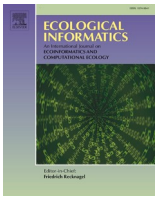
Ecological Informatics 85 (2025) 102995

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf



Improving daily reference evapotranspiration forecasts: Designing AI-enabled recurrent neural networks based long short-term memory

Mumtaz Ali [a,f,*], Jesu Vedha Nayahi [b], Erfan Abdi [c], Mohammad Ali Ghorbani [c,*], Farzan Mohajeri [h], Aitazaz Ahsan Farooque [d,g], Salman Alamery [e]

[a] UniSQ College, University of Southern Queensland 4305 QLD, Australia
[b] Anna University Regional Campus –, Tirunelveli, India,
[c] Department of water engineering, University of Tabriz, Tabriz, Iran
[d] Canadian Centre for Climate Change and Adaptation, University of Prince Edward Island, St Peters, PE, Canada
[e] Department of Biochemistry, College of Science, King Saud University, Saudi Arabia
[f] Scientific Research Center, Al-Ayen University, Thi-Qar, Nasiriyah 64001, Iraq
[g] Faculty of Sustainable Design Engineering, University of Prince Edward Island, Charlottetown, PE, Canada
[h] Islamic Azad University, Central Tehran Branch, Iran

ARTICLE INFO

Keywords:
Agriculture
Deep learning
ETo
Forecasting
RNN-LSTM
CNN-LSTM

ABSTRACT

Predicting daily reference evapotranspiration (ETo) plays a significant role in numerous environmental and agricultural applications. It aids in optimizing agricultural practices, enhancing drought resilience, supporting environmental conservation efforts, and providing critical data for research. By leveraging advanced technologies and accurate modeling techniques, stakeholders can make informed decisions that promote sustainability and resilience in the face of changing climatic conditions. The main purpose of this investigation was to forecast the daily ETo trends at Melbourne and Sydney stations in Australia, where several cutting-edge machine learning methodologies were employed. The modeling approach encompassed the implementation of Neural Network (NN), Deep Learning (DL), Recurrent Neural Networks (RNN), RNN based Long Short-Term Memory (RNN-LSTM), and Convolutional Neural Network based LSTM (CNN-LSTM) to forecast daily ETo using historical meteorology data. During the model development stage, the optimal variables were determined successfully via heatmaps for precise assessment of ETo in both stations. The predictive models were built by incorporating both the training subset (80 %, covering the years 2009 to 2020) and the testing subset (20 %, ranging from 2021 to 2024) independently to forecast ETo. The results confirmed that the RNN-LSTM attained higher prediction accuracy as compared to NN, DL, RNN, and CNN-LSTM models. Conversely, based on the visual representations and assessments, one can grasp the significant resemblance between the forecasts of the RNN-LSTM model and the actual data. By combining RNNs with LSTM units, models can leverage the strengths of both approaches to improve their ability to process sequential data effectively. This integration allows for better capturing of both short-term and long-term dependencies in the input sequences. Upon careful evaluation, it became clear that the error values associated with the RNN-LSTM models were negligible at the designated stations during the testing phase, with an RMSE of 0.0011 mm for Melbourne, and 0.022 mm for Sydney, followed by RNN, DL, and NN respectively. The proposed modeling approach can be beneficial in monitoring and managing water and crop planning which relies on precise ETo predictions.

1. Introduction

Time series analysis employs a statistical approach to investigate regularly collected data points to uncover the underlying patterns and trends. Identifying patterns, forecasting future data points, and making informed decisions rely on temporal relationships among variables. The arrangement of time series data is happening chronologically to facilitate the examination of patterns such as trends, cycles, and seasonal

* Corresponding authors at: UniSQ College, University of Southern Queensland, 4305, QLD, Australia.
E-mail addresses: mumtaz.ali@unisq.edu.au (M. Ali), jesuvedhanayahi.j@auttvl.ac.in (J.V. Nayahi), Ghorbani@tabrizu.ac.ir (M.A. Ghorbani), afarooque@upei.ca (A.A. Farooque), salamery@ksu.edu.sa (S. Alamery).

https://doi.org/10.1016/j.ecoinf.2025.102995
Received 17 October 2024; Received in revised form 2 January 2025; Accepted 2 January 2025
Available online 3 January 2025

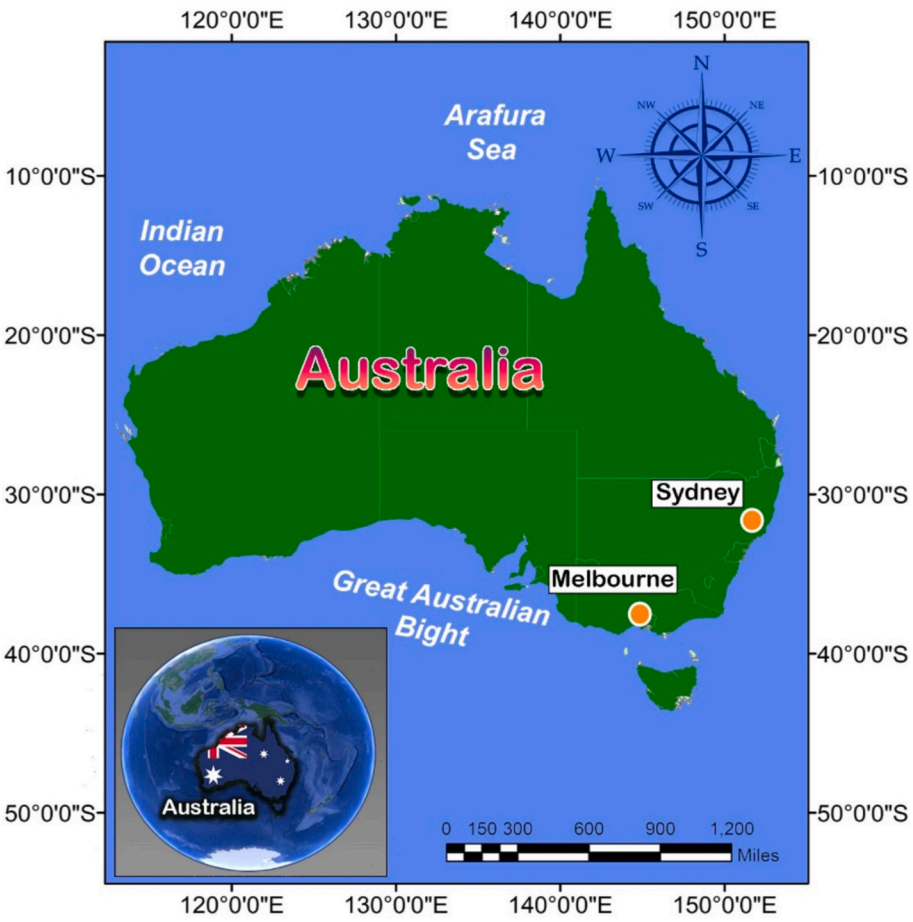


Fig. 1. Location map of Melbourne and Sydney in Australia.

differences to predict upcoming occurrences (Kirchgässner et al., 2012). Time series forecasting is pivotal in various key areas of contemporary agriculture, such as scheduling irrigation, modeling crops, and managing agricultural water (Richetti et al., 2023).

ETo plays a vital role in the hydrologic cycle as it interlinks the cycles of water (evaporation), energy (latent heat flux), and carbon (transpiration-photosynthesis balance) (Fisher et al., 2017). ETo is the sum of water evaporation and transpiration from the Earth's surface into the atmosphere. Hence, the transfer of water from the Earth's surface to the atmosphere, involving evaporation from soil, water bodies, and vegetated canopies, as well as transpiration from plants, is known as ETo. The Earth's biosphere, hydrosphere, and atmosphere rely on the ETo to maintain a stable water and energy equilibrium, particularly in areas with scarce water resources where it influences nearly the entirety of the surface water budget (Morillas et al., 2013; Oki and Kanae, 2006). The assumption of a hypothetical reference crop surface and the application of standardized conditions result in plant surface controlling effects, transforming evapotranspiration into a purely meteorological variable (Allen et al., 1998).

The prediction of the ETo rate becomes possible when one applies the principle of energy conservation and takes into consideration factors like net radiation, sensible heat, soil heat flux, latent heat flux, etc. Recent advancements in Machine Learning (ML) have opened an exciting opportunity to develop data-driven models of physical processes (Ali et al., 2021; Ghimire et al., 2022; Huang, 2009; Lary et al., 2016; Schultz and Wieland, 1997). By accurately estimating ETo, these models have proven their predictive potential and ability to capture complex relationships between input data and output. Different ML algorithms, including Multi-Layer Perceptron (MLP), Random Forest (RF), Artificial Neural Network, and Support Vector Machine (SVM) have been evaluated in terms of their efficacy in predicting daily ETo based on research findings (Fan et al., 2018; Ferreira et al., 2019; Schultz et al., 2000; Sharafi et al., 2024; Talib et al., 2021). Deep Learning (DL) has already had a substantial influence on weather and hydrologic applications compared to other ML methods. The utilization of supervised methods dominates the field of learning and evapotranspiration, with prevalent techniques including Deep Neural Networks or DL, CNN, RNN, and LSTM methods (Chen et al., 2024; Granata and Di Nunno, 2021; Jung et al., 2022; Liu et al., 2024; Roy, 2021).

Using combined models with LSTM offers several advantages over using independent models. This integration can lead to more accurate predictions and a better understanding of the relationships between different types of data (Gelete and Yaseen, 2024). Additionally, combined models can leverage the strengths of LSTM, such as its ability to capture long-term dependencies in sequential data, while also utilizing the strengths of other models to handle specific data types. This can result in improved performance and generalization capabilities. Overall, using combined models with LSTM can lead to more robust and effective machine learning models for a wide range of applications (Wang et al., 2024). Therefore, in this study, three independent models (NN, DL, RNN) and two hybrid models (RNN-LSTM, CNN-LSTM) have been used to evaluate the accuracy and performance of the reference daily evapotranspiration forecast.

The main objective of this study is to investigate the accuracy of novel DL models in predicting daily ETo at two different stations in Australia, by harnessing the strengths of NN, DL, RNN, RNN-LSTM, and CNN-LSTM models. Up until this point, there have been no documented investigations concentrating on predicting the variations in ETo at these two stations using these approaches by analyzing the daily meteorology data. This research is focused on effectively forecasting the daily trends

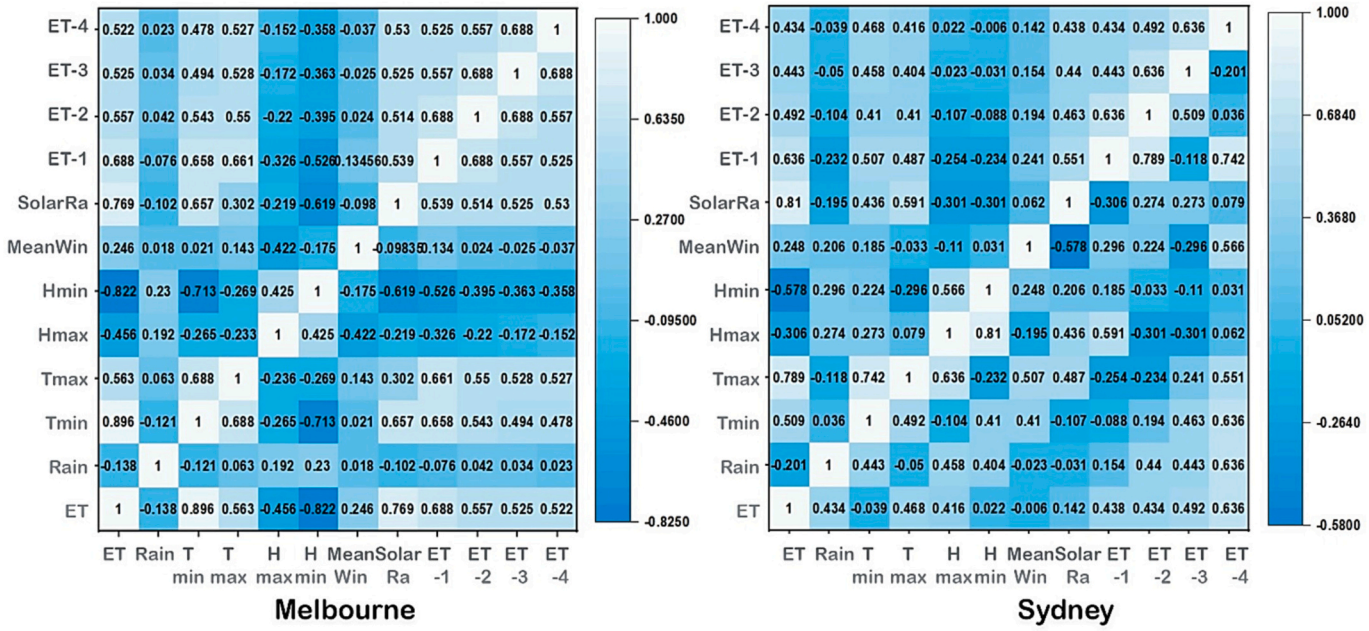


Fig. 2. Heatmap based on the correlation between input parameters and ETo.

in ETo at two Australian locations, namely Melbourne and Sydney, utilizing NN, DL, RNN, RNN-LSTM, and CNN-LSTM models.

2. Materials and methods

2.1. Study area and dataset

The study focused on two important stations, Melbourne, and Sydney in Australia, for which daily datasets of input predictors and targets were acquired from the Bureau of Meteorology (BMA) during the years 2009 to 2024. Fig. 1 shows the location of these two stations. The accurate prediction of daily reference evapotranspiration (ETo) is essential for effective irrigation scheduling, water resource management, and crop growth simulation in Melbourne and Sydney, as well as within the broader Australian agricultural context.

Melbourne is the capital city of the Australian state of Victoria and is located on the southeastern coast of Australia, specifically at the head of Port Phillip Bay. The city's coordinates are approximately 37.8136° S and 144.9631° E. With a population exceeding 5 million individuals, the city encompasses an urbanized territory of approximately 9992 km^2 and holds the distinction of being Australia's second-most inhabited city. The urban area experiences a temperate coastal climate, characterized by balmy summers and brisk winters.

Situated on the southeastern coast of Australia, Sydney claims the title of both its largest and most populous city, overlooking the Tasman Sea. The city's coordinates are approximately 33.8688° S, and 151.2093° E. With a population exceeding 5 million individuals, the city encompasses an urbanized territory of approximately 12,368 km^2. In Sydney, the weather is generally sunny with mild winters and warm summers.

Accurate prediction of ET_o relies on various meteorological parameters, which can vary in importance depending on the model used. In this research, datasets from Melbourne and Sydney stations were used to evaluate the proposed models using four delay reference evapotranspiration (ET_{t-4}), rain, min/max temperature ($T_{min/max}$), max/min humidity ($H_{min/max}$), mean wind speed (Mean Win), and solar radiation time series data with high and low variations. The heatmap based on the correlation of the between the input data and ETo is shown in Fig. 2.

Data preprocessing is an essential step in predicting and modeling ETo to ensure accurate results. For this purpose, missing or inconsistent values were first cleaned and filtered by filling with average value of the relevant input, which prevents the deviation of predictions to a large extent. In the next step, data normalization was done to scale the features in the standard range, which helped to improve the performance of the predictive model. Then feature selection was performed to identify the most relevant variables (rainfall, temperature, humidity, reference evapotranspiration delay) that have a direct effect on ETo. Finally to create the model, the input datasets split into two sections: one comprising 80 % used for training (ranging from the years 2009 to 2020) and the other with 20 % designated for testing purposes (encompassing the timeline of 2021 to 2024) following (Cannas et al., 2006; Fijani et al., 2019; Quilty and Adamowski, 2018). Moreover, the cross-validation or any random sampling technique cannot be adopted here as the time-series data by definition occur in a temporal order/sequence and this order or sequence must be preserved to retain the structure of the series

Table 1
General information about the models.

Model	Type of hyper-parameters	Values
NN, DL, RNN, RNN-LSTM, CNN-LSTM	Network Type	Feed-forward propagation
	Data Division	Train (80 %) Test (20 %)
	Number of Hidden layers (Neurons)	10–20
	Epoch number	100
	learning Function	0.001–0.007
	Activation function	Relu, Tanh, Tan, Sigmoid
	Training function	Adam

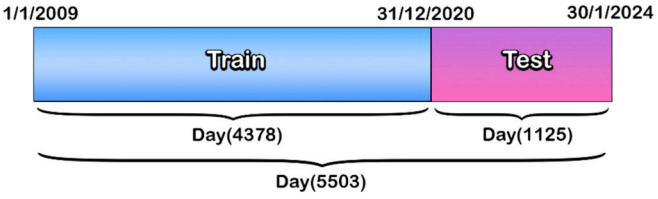


Fig. 3. Graphical train and test.

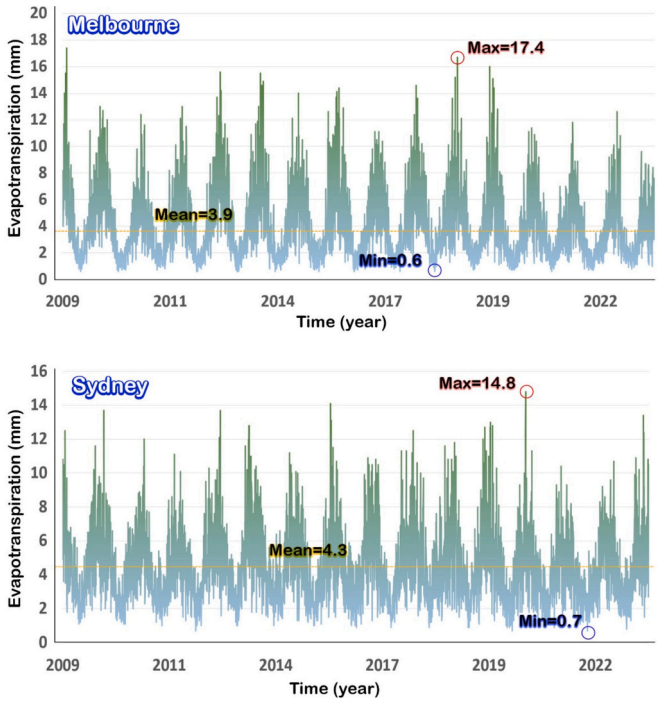


Fig. 4. Daily ETo of the two stations during 2009–2024.

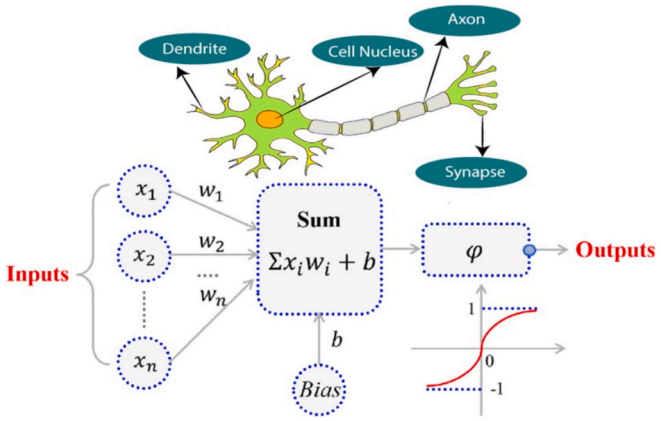


Fig. 5. Neural Network model structure and human brain's biological.

Table 2
Hyper parameters of NN.

Parameter	Value
Linear Layer	8
Dense Unit	1
Epochs	18
Learning Rate	0.01
Activation Function	Tanh, Tan

intact (Bergmeir and Benítez, 2012). Table 1 contains information on the layers and parameters employed in the models that were selected.

Fig. 3 shows the train-test split graphically whereas in Fig. 4, the statistical measures of daily evapotranspiration for the two stations chosen for this investigation are displayed.

This study makes use of a computing platform equipped with the Windows 11 operating system. In order to test and confirm the NN, DL, RNN, CNN-LSTM, and RNN-LSTM models put forth in this research, we conducted trials using the Mathematica 13.3 platform. The experiments are executed in an environment with a NVIDIA GeForce RTX 3050 GPU, an Intel Ryzen 7–6800 CPU, and 16GB RAM.

2.2. Neural Network (NN)

Neural Network (NN), a machine learning model, is categorized under artificial intelligence. The basis of NNs lies in simple mathematical models of the brain, making them a distinct machine learning algorithm. Their usefulness extends to a diverse set of prediction tasks due to their capacity to comprehend intricate nonlinear connections between the response variable and its predictors (Elsner and Tsonis, 1992). In 1943, McCulloch and Pitt's introduction of simplified neurons ignited the fascination with neural networks, resulting in the development of artificial neural networks to generalize mathematical models of biological nervous systems (Abraham. 2005). NNs can adapt to fluctuating data patterns and assimilate fresh information, thereby making them exceptionally suitable for dynamic forecasting environments and capable of effectively dealing with noisy or incomplete data in comparison to traditional models, leading to more dependable forecasts. The distinction between artificial neural networks (ANNs) and deep learning lies in their varying complexity and structure. Various essential parts are interconnected within neural networks to process information and build knowledge from datasets. Understanding these key components is vital for gaining insight into the operations of neural networks (Goyal and Parashar, 2018). Within a neural network, the first layer serves as the input layer and connects with neurons that are specific to distinct features present in the dataset. Neurons play a crucial role as the primary elements in a neural network. Each individual neuron takes in various inputs, gives them specific weights, calculates the sum of all inputs, and processes it through an activation function to produce its final output. The strength of neural connections is determined by factors termed weights (Narayanan and Menneer, 2000).

$$Z = \sum_{i=1}^{n} w_i x_i + b \tag{1}$$

where *wi* are the weights, *xi* are the inputs, and *b* is the bias term.

The adjustment of connection weights between neurons during training is aimed at reducing prediction errors, and biases serve as additional parameters that aid in better data fitting by adjusting the activation function. Introducing activation functions into the model allows for the exploration and learning of intricate patterns, adding a nonlinear aspect. Fig. 5 shows the structure of neural network model inspired by the human brain's biological function with different parameters. Table 2 shows the structural specifications along with parameters of NN.

2.3. Deep Learning (DL)

Categorized under supervised machine learning models, Deep Learning (DL) stands within the domain of artificial intelligence and possesses immense significance by effectively addressing a wide range of complex or seemingly impossible problems that conventional algorithms or human professionals encounter difficulties with. DL carries out the reception, processing, and generation of information similarly to biological neuronal networks (LeCun et al., 2015). DL algorithms are extensively utilized by industries involved in solving complex challenges like meteorology and climate change as they train machines through example-based learning (Abiodun et al., 2018). Generally, each neuron computes a weighted sum of its inputs, applies an activation function, and passes the result to the next layer.

$$y = f(wx + b) \tag{2}$$

f is the activation function (e.g., $ReLU : f(x) = max(0, x)$, $sigmoid : f(x) = \frac{1}{1+e^{-x}}$) that introduces non-linearity into the mode. These functions determine whether a neuron should be activated or not, adding complexity to the model.

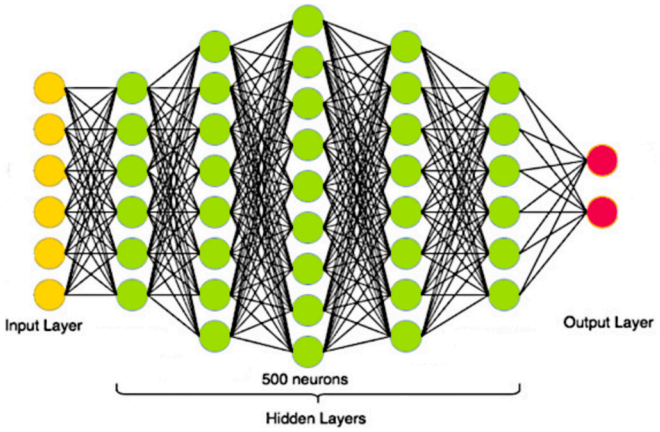


Fig. 6. Deep Learning model structure.

Table 3
Hyper parameters of DL.

Hyper parameters	Value
Linear Layer	12
Batch Normalization Layer	3
Softmax Layer	1
Dense Unit	1
Epochs	38
Learning Rate	0.01
Activation Function	Tanh, Tan, Sigmoid

By automatically learning features from the data, DL algorithms eliminate the requirement for hand-engineered features. Traditional machine learning methods may struggle with handling large and complex datasets, but DL algorithms are up to the task. The capability of deep learning models to detect and assimilate non-linear relationships in complex datasets makes them highly valuable for data analysis (Dong et al., 2021). A common DL model structure is depicted in Fig. 6. Table 3 shows the structural specifications and hyper parameters of DL.

2.4. Recurrent neural network (RNN) and long short-term memory (LSTM)

RNN is one of the most suitable DL models for time series or temporal data due to its memory component. This memory helps to remember the learning from the previous sequence and applies that knowledge to the next sequence in the input. As RNN is a cyclic network and the output at time sequence t_n is used to predict the output at time sequence t_{n+1}. In our real time experiment, the ETo data is a sequential data collected on a daily basis and hence RNN is a promising method for predicting the ETo values. The RNN works on the same principle with few changes in CNN. NN uses each unit in the network, but RNN uses the same weight for all the units in the same layer. The feedback loop in the architecture is another significant change. The output from the previous step is given as input to the current step. This helps the network to make improved predictions based on the previous output. The hidden state act as the memory of the RNN which uses different activation functions such as sigmoid, tanh etc. Although RNN is more suitable for sequential or time series data, it has two major limitations termed as vanishing gradient and exploding gradients (Ye et al., 2019).

Assume the weight in the first-time step is greater than 1. The weights in each network of a particular time step are given as input to the next time step and this is multiplied with the next input. This step gets repeated as many number of times as the number of time steps. This will cause the gradients to have a very high value and this is termed as exploding gradients. If the weight is less than 1, multiplying it with the input will result in gradients close to zero, which is termed as vanishing gradients. An RNN processes sequences of inputs by maintaining a hidden state that captures information from previous time steps. The basic equations governing an RNN can be expressed as follows.

$$h_t = f(w_h h_{t-1} + w_x x_t + b) \tag{3}$$

Where h_t is the hidden state at time t_b f is a non-linear activation function (often tanh or ReLU), W_h is the weight matrix for the hidden state, W_x is the weight matrix for the input, x_t is the input at time t_t and b is the bias term.

In order to overcome the exploding gradients and vanishing gradients, a variant of RNN is proposed known as Long Short-Term Memory (LSTM). They achieve this through a gating mechanism that controls the flow of information. This memory cell is controlled by three gates (Sherstinsky and Alex, 2020). The input gate determines which new information to store in the memory cell from both the current input and previous hidden state. The Forget Gate determines which details will be eliminated from the previous memory cell state. The Output Gate determines which information should be output from the present input, past hidden state, and memory cell.

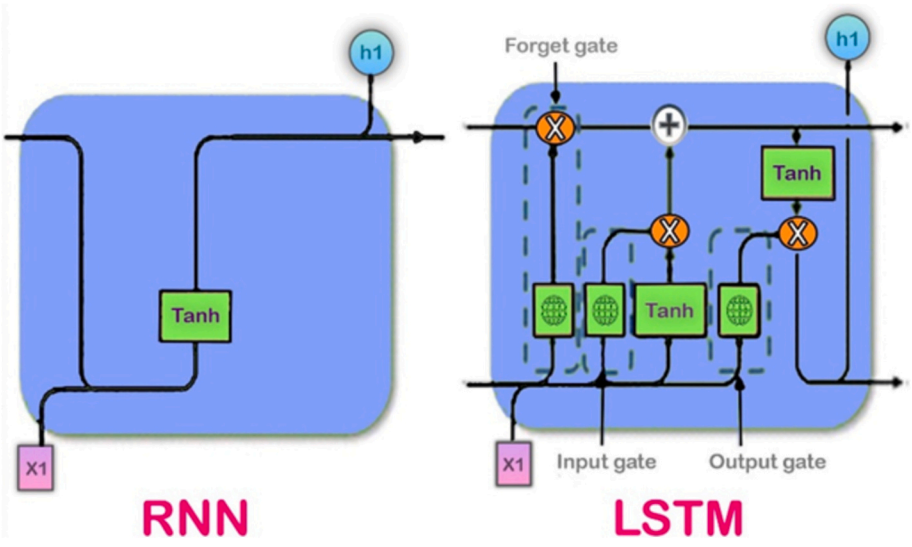


Fig. 7. The structure of recurrent neural network model and short-term memory with tanh, sigmoid activation layers.

Table 4
Hyper parameters of RNN-LSTM.

Hyper parameters	Value
Simple RNN Layer Units	32
LSTM Units	10
Dense Unit	1
Epochs	100
Learning Rate	0.01
Activation Function	Sigmoid, Relu

$$p_s = \sigma\left(W_{p[h_{s-1},x_s]} + b_p\right) \quad (4)$$

$$q_s = \sigma\left(W_{q[h_s-1,x_s]} + b_s\right) \quad (5)$$

At time s, the prior hidden state is denoted as h_{s-1}, while the weight matrix is represented by W_p and the input is denoted as x_s; additionally, the bias vector is b_p and the forget gate activation vector is simply referred to as p.

$$h\,s = f_s \times tanh(v_s) \quad (6)$$

RNN network is modified to include both long term memory and short-term memory (da Silva and de Moura Meneses, 2023). Fig. 7 shows the structure of this model in detail and Table 4 shows the structural specifications of RNN-LSTM.

Combining Recurrent Neural Networks (RNNs) with Long Short-Term Memory (LSTM) models offers several advantages, particularly in the realm of sequential data processing. LSTMs are specifically designed to address the limitations of standard RNNs, particularly their inability to retain information over long sequences due to the vanishing gradient problem. By incorporating memory cells and gating mechanisms, LSTMs can effectively manage both long-term and short-term dependencies, making them superior for tasks where context from earlier inputs is crucial (Al-Selwi et al., 2024). Hence, combining RNNs with LSTM models significantly enhances the ability to process sequential data by improving memory retention, accuracy, gradient handling, versatility across applications, and managing complexity effectively.

2.5. Performance criteria

Evaluation of the effectiveness of modeling during the test stages includes the analysis of correlation coefficient (R), coefficient of determination (R^2), and root mean square error (RMSE). These metrics not only enhance model validation and predictive accuracy but also play a crucial role in effective water management strategies, particularly in regions facing water challenges like Australia. By leveraging these statistical tools, stakeholders can make informed decisions that promote sustainability and resilience in agricultural practices. The correlation coefficient measures the strength and direction of a linear relationship between two variables and the spectrum spans from −1 to 1, with 1 denoting a perfect positive linear connection, 0 representing no linear connection, and −1 representing a perfect negative linear connection (Schuurmann et al., 2008). The coefficient of determination represents the proportion of the total variation in the output variable that is accounted for by the fitted model. This coefficient lies within the interval $0 \le R^2 \le 1$, and a zero R^2 indicates that the predictive model does not explain the variance of the observed data set. A measure known as RMSE calculates the typical deviation between the predicted values from a statistical model and their corresponding actual values. In mathematical terms, the standard deviation of the residuals reflects the difference between the regression line and the data points whose value is from zero to infinity (Aptula et al., 2005).

$$R = \frac{\sum_{i=1}^{N}(ET_{oi} - \overline{ET_o})(ET_{pi} - \overline{ET_p})}{\sqrt{\sum_{i=1}^{N}(ET_{oi} - \overline{ET_o})^2 \cdot \sum_{i=1}^{N}(ET_{pi} - \overline{ET_p})^2}} - 1 \le R \le 1 \quad (7)$$

$$R^2 = \left(\frac{\sum_{i=1}^{N}(ET_{o,i} - \overline{ET_o})(ET_{p,i} - \overline{ET_p})}{\sqrt{\sum_{i=1}^{N}(ET_{o,i} - \overline{ET_o})^2 \cdot \sum_{i=1}^{N}(ET_{p,i} - \overline{ET_p})^2}}\right)^2 \quad 0 \le R^2 \le 1 \quad (8)$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(ET_{pi} - ET_{oi})^2} \quad 0 \le RMSE \le \infty \quad (9)$$

where, ET_{pi} and ET_{oi} indicate the estimated and observed data, N refers to the number of data, and $\overline{ET_p}$ and $\overline{ET_o}$ are the average of the estimated and observed data.

The time series forecasting techniques are classified in two categories i.e., univariate, or multivariate, depending on the number of variables to be used. In this study assuming that the future values of the ETo time series depend on the past values of evapotranspiration and as well as precipitation, minimum and maximum temperature, minimum and

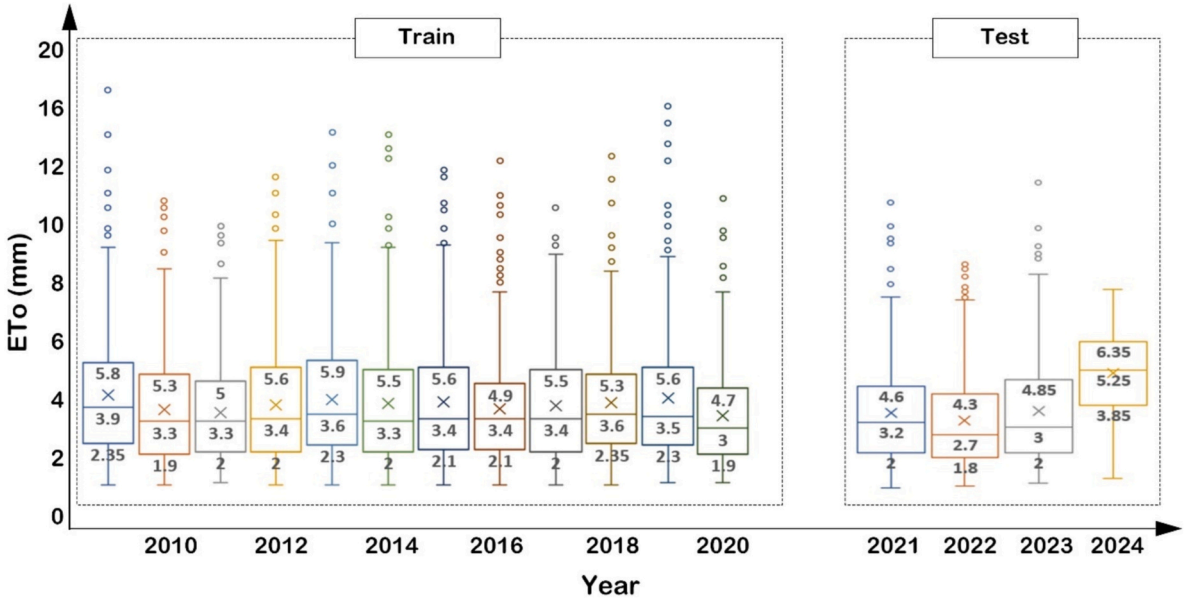


Fig. 8. Boxplot of the ETo data for Melbourne city with statistical profile.

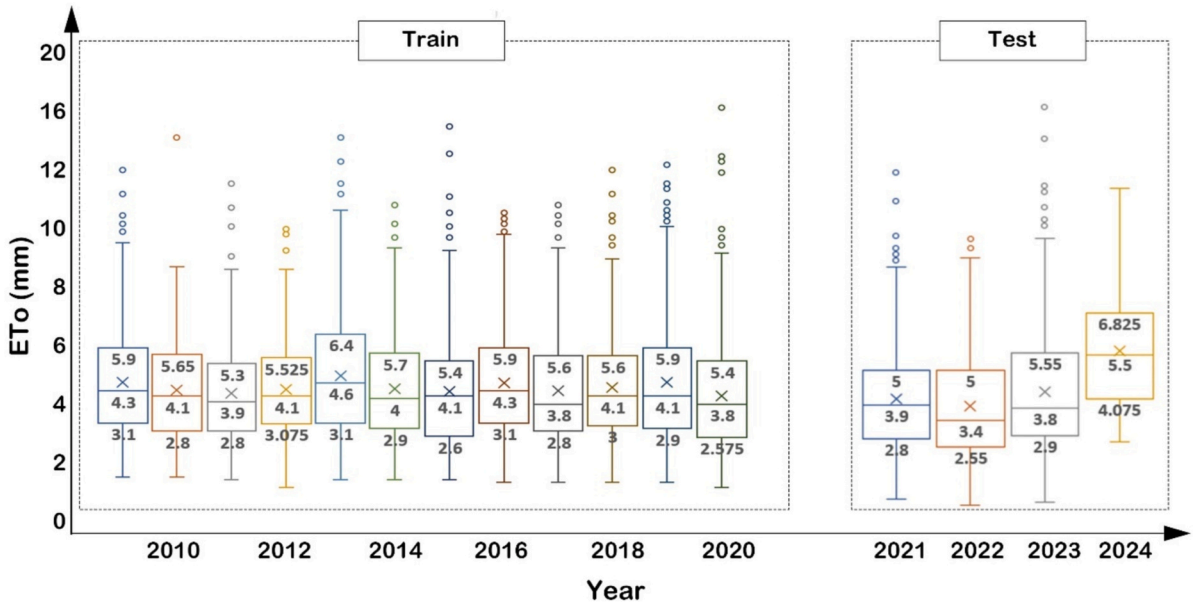


Fig. 9. Box plot of the ETo data for Sydney city with statistical profile.

Table 5
Performance criteria of the DL and NN models in the testing periods for two stations.

Cities		Testing (ET)		
	Model	R	R^2	RMSE (mm)
Melbourne	NN	0.9980	0.8720	0.71
	DL	0.9990	0.9960	0.122
	RNN	0.9998	0.9990	0.051
	CNN-LSTM	0.9997	0.9995	0.0434
	RNN-LSTM	**0.9999**	**0.9999**	**0.0011**
Sydney	NN	0.9090	0.8262	1.026
	DL	0.9920	0.9820	0.253
	RNN	0.9999	0.9998	0.023
	CNN-LSTM	0.9995	0.9980	0.074
	RNN-LSTM	**0.9999**	**0.9999**	**0.022**

maximum relative humidity, wind speed, and solar radiation. Therefore, the multivariate forecasting method for prediction of future ETo was used here as this method has some advantages over univariate time series forecasting method.

3. Result

Statistical and visual analysis were employed to evaluate the effectiveness of DL, NN, RNN, CNN-LSTM, and RNN-LSTM models in predicting daily ETo.

Fig. 8 and Fig. 9 shows the distribution of the data collected from the two cities Sydney and Melbourne respectively from years 2009 to 2024. Each box with whiskers represents the distribution of ETo values in each year. The plot shows that there are outliers in training and test datasets collected from both the stations.

The performance measures such as the correlation coefficient (R), coefficient of determination (R^2), and root mean square error (RMSE),

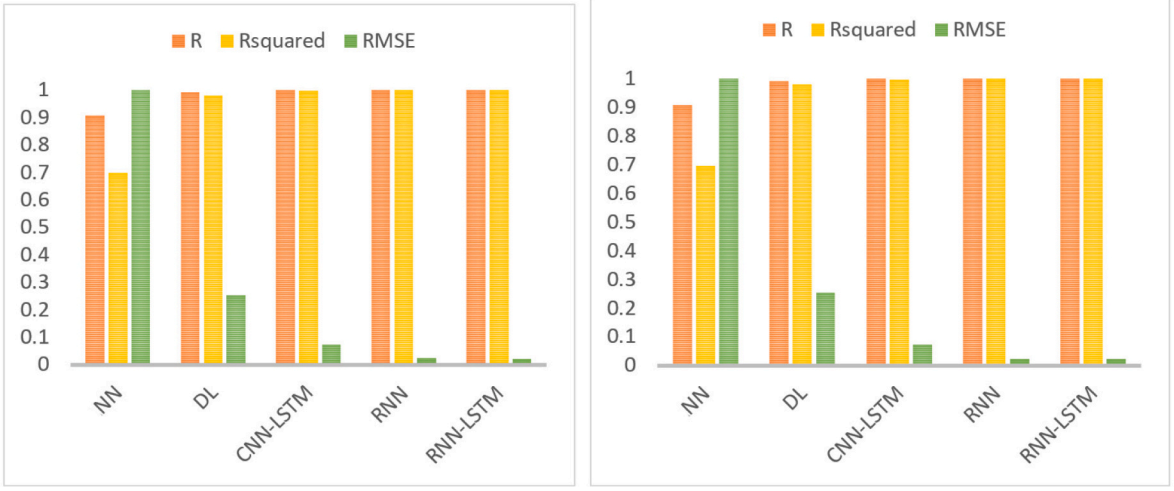


Fig. 10. Performance comparison of the models in terms of R, RSquared and RMSE for (a) Melbourne, and (b) Sydney.

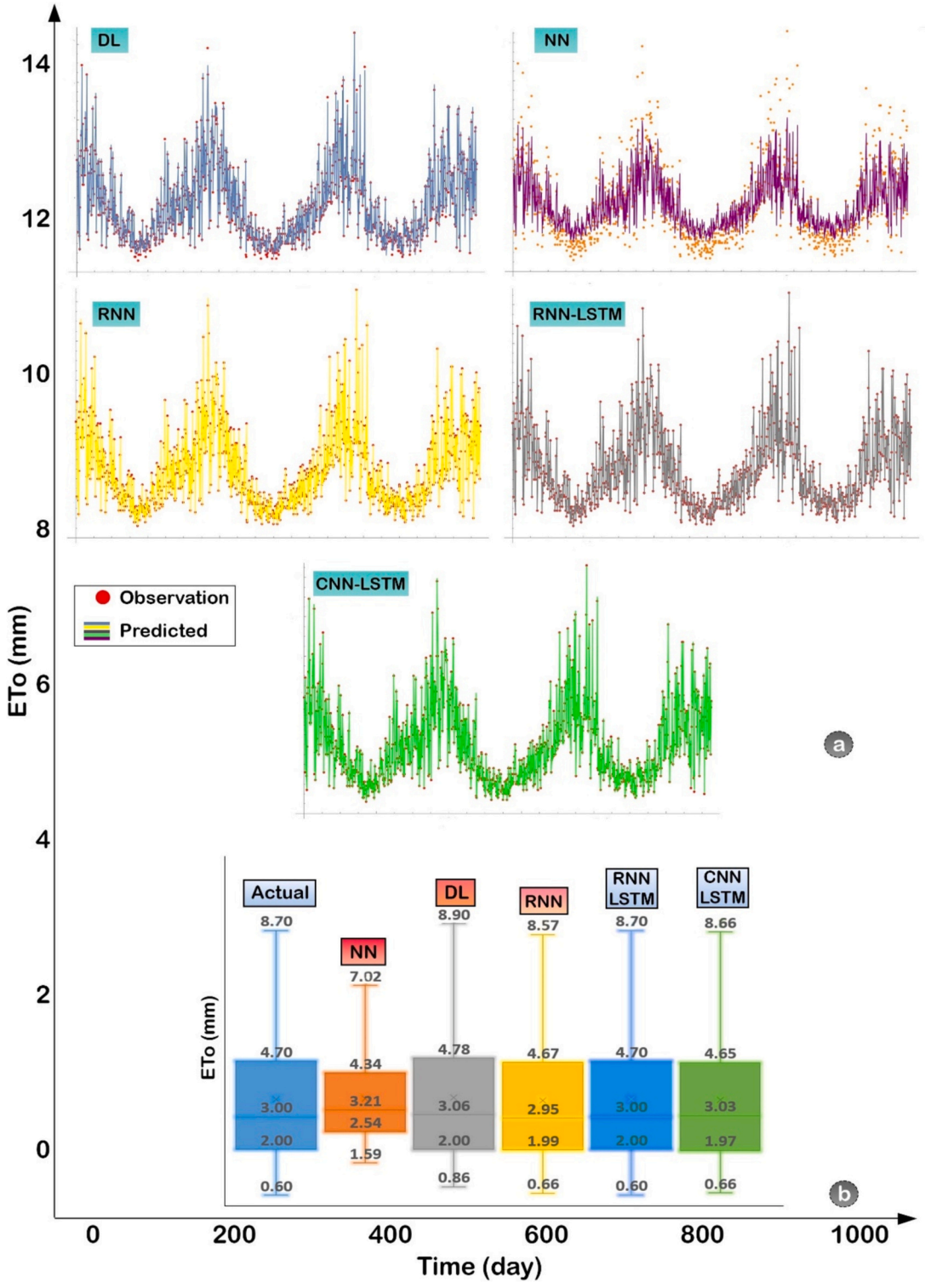


Fig. 11. Time series plot, and box plot of the observed and forecasted ETo of the models for Melbourne. The small circles are the daily ETo data in (a) whereas the vertical and horizontal lines in (b) stand for the five number summary (i.e., minimum, first quartile, median, third quartile, and maximum).

have been computed for the test periods and are listed in Table 5 and Fig. 10 shows the evaluation of the models graphically.

The RNN-LSTM method outperforms other models in terms of all three evaluation criteria for the cities of Melbourne and Sydney. Hence, exhibits a RMSE value in the range of $0.0253 < \text{RMSE} < 0.0011$, correlation coefficient R in the range of $0.992 < R < 0.9999$, and R^2 value in the range of $0.982 < R^2 < 0.9999$ for both stations. It is also to be noted that CNN-LSTM appeared to be the 2nd best predictive model for Melbourne city whereas RNN achieved the 2nd ranking for Sydney based on R, R^2, and RMSE value (Table 3 and Fig. 10). Moreover, the performance of these machine learning models is acceptable for both selected regions to predict daily ETo but overall, the RNN-LSTM appeared to be the most accurate model to predict daily ETo for both stations.

Fig. 10. shows the R, Rsquared (i.e., R^2) and the RMSE of the prediction models used in this experiment. The RMSE of the RNN-LSTM

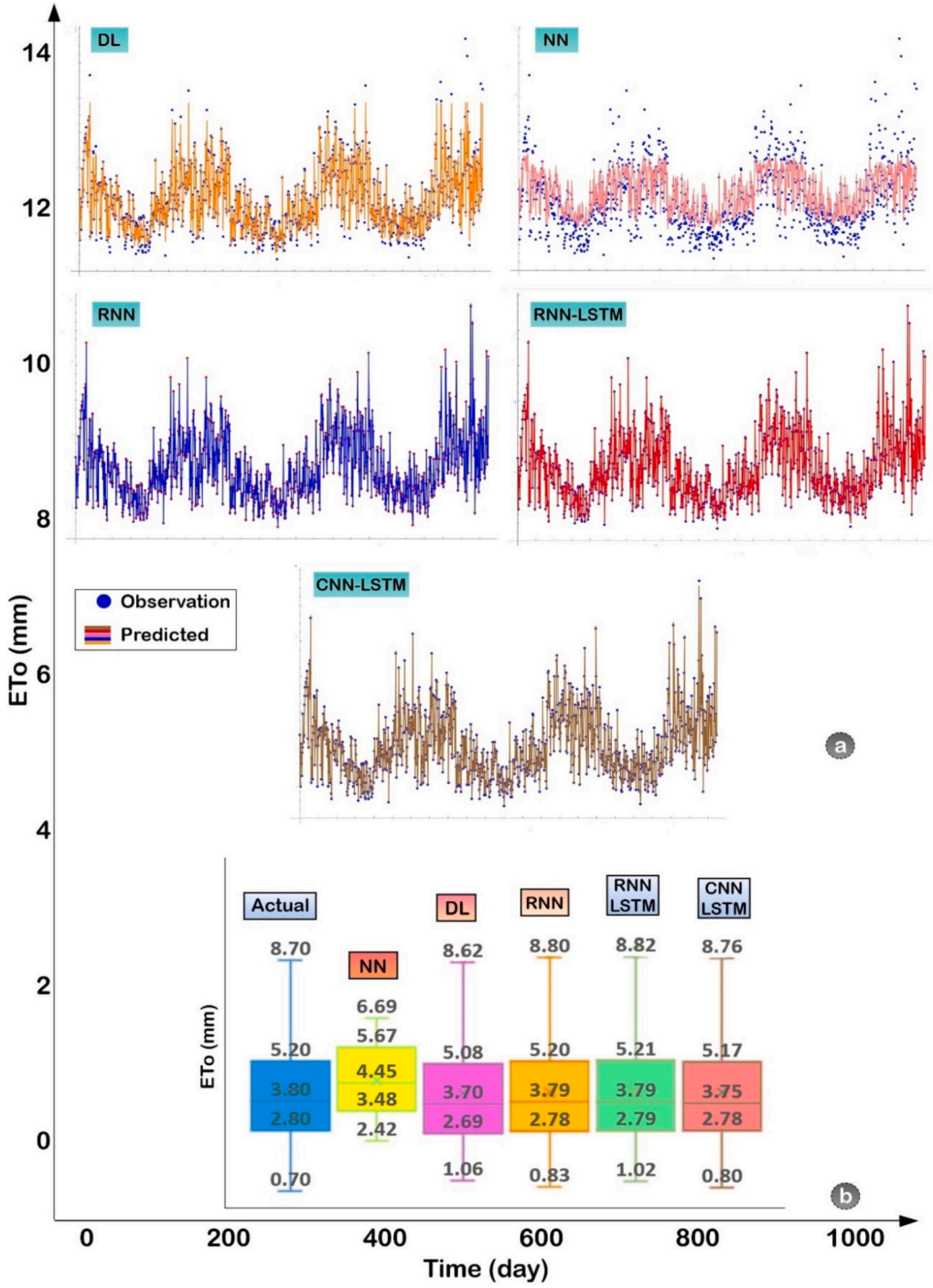


Fig. 12. Time series plot, and box plot of the observed and forecasted ETo of the models for Sydney. The small circles are the daily ETo data in (a) whereas the vertical and horizontal lines in (b) stands for the five number summary (i.e., minimum, first quartile, median, third quartile, and maximum).

model is close to zero and the correlation coefficient and the R squared is almost close to 1 for both Melbourne and Sydney stations. This confirms that RNN-LSTM outperforms other comparing models in both stations to predict daily ETo.

Fig. 11 and Fig. 12 shows the time series plot along with the box plot of the predicted ETo values generated by the machine learning models with respect to the actual ETo recorded on a particular day. Fig. 11a shows the time series plot of the predicted ETo of NN, DL, CNN-LSTM, RNN, RNN-LSTM respectively for Melbourne station. The box plot in Fig. 11b compares the distribution of the actual and predicted ETo using NN, DL, RNN, RNN-LSTM, and CNN-LSTM model in Melbourne station. The distribution of the predicted ETo of RNN-LSTM is much closer to the actual ETo value, following by CNN-LSTM, DL, RNN, and NN models respectively.

Fig. 12a shows the time series plot of the predicted ETo value of the NN, DL, CNN-LSTM, RNN, and RNN-LSTM models for the Sydney station. Furthermore, Fig. 12b showcases a boxplot comparing the actual ETo distribution to the predicted ETo distribution from these models (i.e., NN, DL, RNN, RNN-LSTM, and CNN-LSTM). The distribution of the predicted ETo of RNN- LSTM are much closer to the actual ETo distribution when compared to other models. Moreover, both the RNN and CNN-LSTM models have also an acceptable performance for predicting

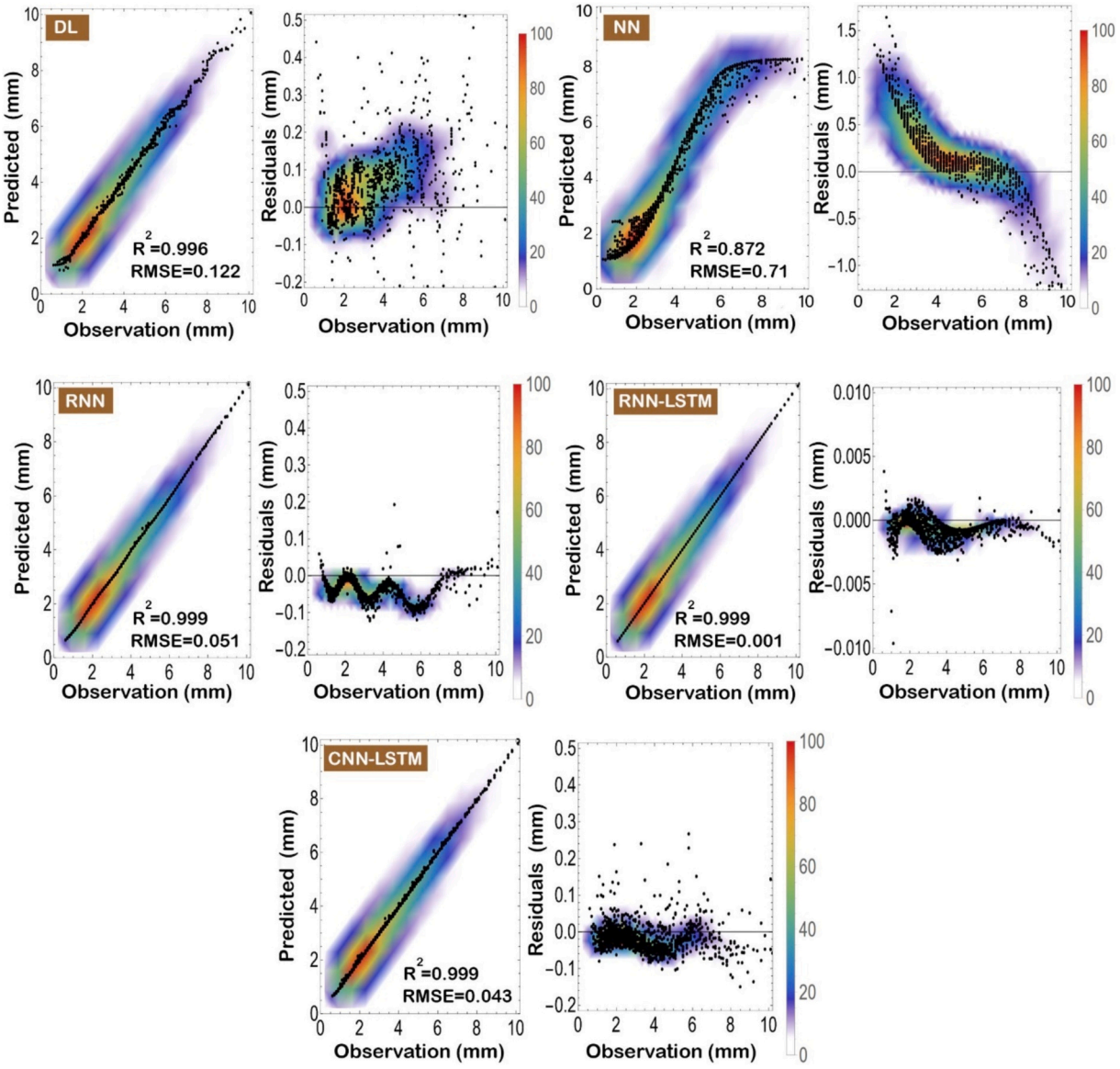


Fig. 13. Residual errors, and scatter plots of the observed and predicted ETo of the five models for Melbourne city in testing period. The colored shadows provide a smooth representation of the distribution of dataset showing the relative frequency.

daily ETo in Sydney station.

Fig. 13 shows the scatter and residual plots of NN, DL, CNN-LSTM, RNN, and RNN-LSTM models for the city of Melbourne. According to this figure, NN and DL models with more dispersion and residual error (− 1 to 1.5) have poor performance. On the other hand, RNN, and CNN-LSTM models with less dispersion and residual error (− 0.2 to 0.5) have performed better than NN and DL models. However, according to the Fig. 13 related to RNN-LSTM model, it can be seen that this model has performed better than other models with minimum dispersion and residual error (− 0.01 to 0.01). Also, Fig. 14 shows the scatter and error residuals for the city of Sydney. It is visible that the RNN-LSTM model with less dispersion and residual error (− 0.04 to 0.04) is the best performing model for predicting daily ETo. On the other hand, RNN model with residual error (− 0.05 to 0.02), CNN-LSTM (− 0.2 to 0.2), DL (−0.8 to 0.4), and NN model (− 1 to 1.5) had acceptable performance, respectively.

4. Discussion

The research successfully attained precise daily predictions of ETo by employing the DL models incorporating complex layers. Utilizing various statistical indicators and graphical representations, it was evident that the RNN-LSTM exhibited superior performance compared to the benchmarking models when predicting ETo in Melbourne and Sydney stations. The inclusion of ensemble learning in the DL methodology leads to greater accuracy and decreased variability, ultimately reducing overfitting and enhancing model stability. Also, according to the correlations in the heat map, the meteorological parameters of minimum temperature, minimum relative humidity, and the amount of solar radiation have a greater effect on ETo than the rest of the inputs. The compatibility and accuracy of this study can be found from previous studies where the adoption of a DL model resulted in a notable upsurge (Babaeian et al., 2022; Chen et al., 2020; Zhang et al., 2023). However, the implementation of DL model along with complex layers incorporating meteorological parameters as inputs to predict ETo in Melbourne and Sydney stations has not yet fully explored.

Weekly predictions of ETo were researched by Karbasi et al. (2022) using an Auto Encode Decoder Bidirectional Long Short-Term Memory (AED-BiLSTM) hybrid deep learning model. Also, they used Generalized Regression Neural Network (GRNN) and Extreme Gradient Boosting

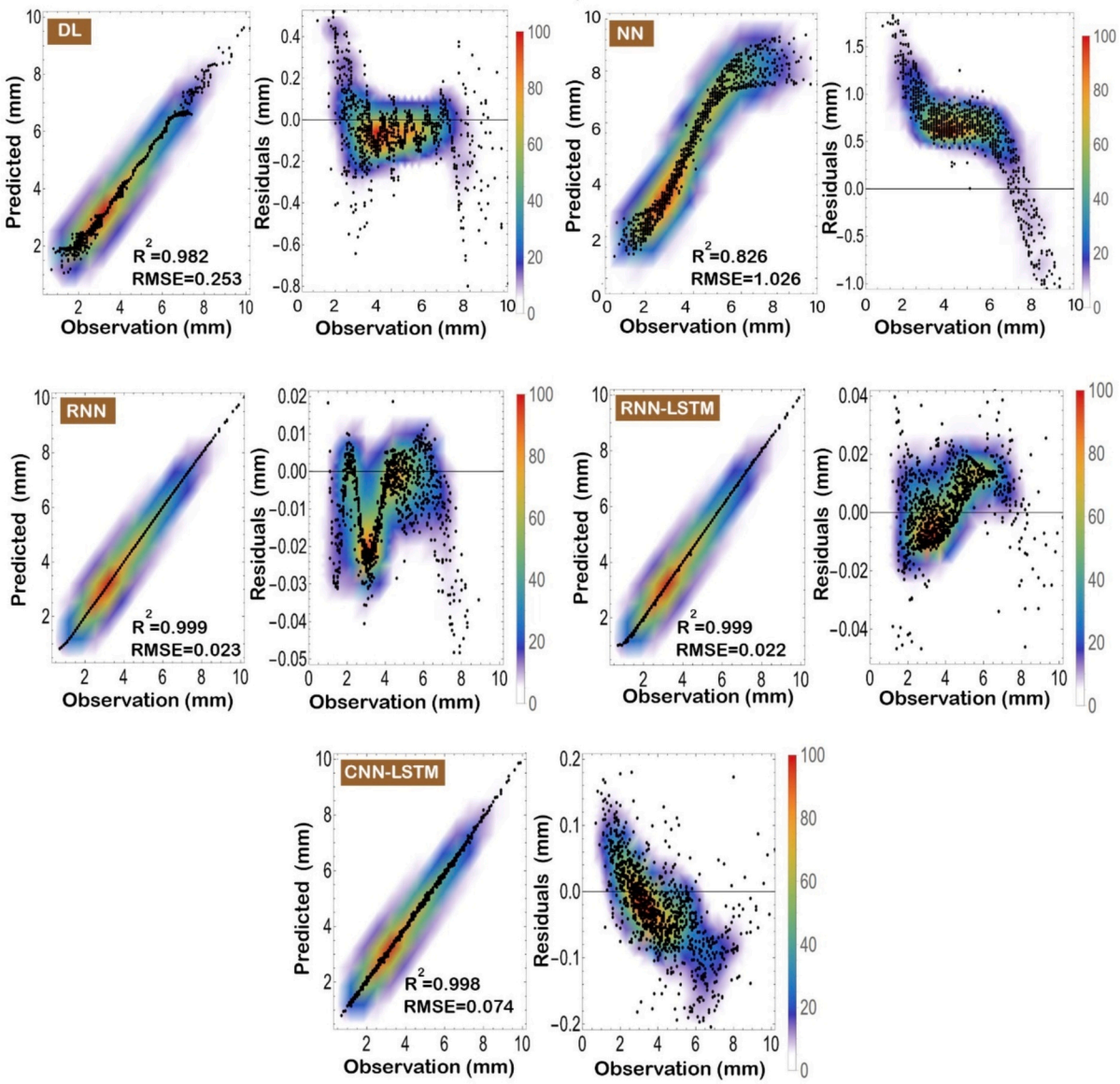


Fig. 14. Residual errors, and scatter plots of the observed and predicted evapotranspiration of the five models for Sydney city in testing period. The colored shadows provide a smooth representation of the distribution of dataset showing the relative frequency.

(XGBoost) machine learning models to evaluate the newly developed model. Compared to the GRNN and XGBoost models, the newly developed AED-BiLSTM model demonstrated a higher level of accuracy in predicting weekly ETo with an R value of 0.965 and an RMSE of 3.21 mm/week. Ferreira and da Cunha (2020) investigated the use of two artificial neural network (ANN) and support vector machine (SVM) models to forecast future ETo values in a more advanced manner. The ANN and SVM models were built using a two-step process: first, identifying groups of weather stations with similar climatic features through K-means clustering, then incorporating past meteorological data into the models as input. The results showed that both models have an acceptable performance (mm/dayRMSE = 0.6 and R2 = 0.81).

Chia et al. (2022) focused on long-term forecasting of monthly reference evapotranspiration (ET0) using deep neural network models. They compared different training strategies and approaches, especially evaluating one-dimensional convolutional neural networks (CNN-1D), short-term memory (LSTM) networks, and gated recurrent unit (GRU) networks at four stations in Peninsular Malaysia. The best-performing GRU models achieved an average mean absolute error (MAE) of 0.182 mm/day, root mean square error (RMSE) of 0.260 mm/day, and a mean percentage error (MAPE) of 4.972 %, with a Kling-Gupta efficiency (KGE) of 0.747.

Vaz et al. (2022) to estimate ETo explored various machine learning algorithms including Artificial Neural Networks (ANN), Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), Recurrent Neural Network (RNN), and hybrid neural network models such as LSTM-ANN, RNN-ANN, and GRU-ANN. The LSTM-ANN model with a determination coefficient of 0.967 and an error of 0.307 mm/day had a high performance compared to other models. Saggi and Jain (2019) utilizing the Multilayer Perceptron (MLP), the daily ETo for Hoshiarpur and Patiala Districts in Punjab was computed. Compared to machine learning algorithms like Random Forest (RF), Generalized Linear Models (GLM), and Gradient Boosting Machine (GBM), the MLP showcased better performance.

CNN-LSTM models are particularly effective for long-term predictions due to their ability to capture spatial features through the convolutional layers before processing temporal sequences with LSTM layers. This dual capability allows them to maintain high accuracy over extended forecasting horizons. In contrast, RNN-LSTM models may excel in scenarios requiring immediate temporal predictions but can struggle

with longer sequences due to issues like vanishing gradients, which LSTMs are designed to mitigate but may still face in certain configurations. Overall, meta-heuristic-driven ML models have demonstrated their capacity to achieve notable successes and showcase immense potential. Also, according to the conducted studies, studies in which LSTM models have been combined have provided more accurate predictions, and in this research, acceptable results were obtained by combining RNN and LSTM.

5. Summary and conclusion

Accurate predictions of nonlinear processes, particularly ETo which is influenced by rainfall, temperature, and sunshine, have been significantly improved through the utilization of ML techniques. The fundamental aim of this research was to determine the effectiveness of five distinct ML techniques, namely DL, NN, RNN, CNN-LSTM and RNN-LSTM in forecasting daily ETo for prominent Australian urban cities Melbourne and Sydney. The evaluation of these constructed models involved considering the statistical indicators R, R^2, and RMSE, in addition to analyzing the statistical graphs. By producing exceptional outcomes for two selected stations in predicting ETo, the RNN-LSTM model exemplified its dominance over the NN and other comparing models and substantiated its superiority.

One of the primary challenges in applying RNN-LSTM hybrid models is the effective initialization and optimization of the network parameters. Poor initialization can lead to suboptimal learning outcomes, while the complexity of tuning hyperparameters can hinder the model's overall accuracy. Despite LSTMs being designed to mitigate the vanishing gradient problem inherent in vanilla RNNs, they still face difficulties in learning long-term dependencies effectively. The vanishing gradient issue can still arise during training, particularly with very deep networks or when dealing with long sequences, leading to challenges in capturing essential temporal patterns in transpiration data. In the pursuit of enhancing the organization of self-sufficient intelligent models, future investigations should prioritize incorporating DL and complementary hybrid techniques to forecast additional variables (e.g., cloud cover and Human Factors), while concurrently conducting assessments against them.

CRediT authorship contribution statement

Mumtaz Ali: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Conceptualization. **Jesu Vedha Nayahi:** Writing – original draft, Visualization, Methodology, Investigation, Conceptualization. **Erfan Abdi:** Writing – original draft, Validation, Methodology, Conceptualization. **Mohammad Ali Ghorbani:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Farzan Mohajeri:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Aitazaz Ahsan Farooque:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Salman Alamery:** Writing – review & editing, Methodology, Conceptualization.

Acknowledgements

The authors would also like to state that there is no conflict of interest. The authors extend their appreciations for funding provided by the Researchers Supporting Project number (RSP2025R241), King Saud University, Riyadh, Saudi Arabia. The authors are thankful to the Bureau of Meteorology, Australia for providing the required datasets.

Data availability

The data can be access in the following links http://www.bom.gov.au/watl/eto/. The codes can be accessed using the link https://github.com/aiclimatemodeller/My-codes.git.

References

Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A., Arshad, H., 2018. State-of-the-art in artificial neural network applications: a survey. Heliyon 4 (11).
Ali, M., Prasad, R., Xiang, Y., Khan, M., Farooque, A.A., Zong, T., Yaseen, Z.M., 2021. Variational mode decomposition based random forest model for solar radiation forecasting: new emerging machine learning technology. Energy Rep. 7, 6700–6717.
Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-guidelines for computing crop water requirements-FAO irrigation and drainage paper 56. Fao, Rome 300 (9), D05109.
Al-Selwi, S.M., Hassan, M.F., Abdulkadir, S.J., Muneer, A., Sumiea, E.H., Alqushaibi, A., Ragab, M.G., 2024. RNN-LSTM: from applications to modeling techniques and beyond—systematic review. J. King Saud Univ. Comput. Inform. Sci. 102068.
Aptula, A.O., Jeliazkova, N.G., Schultz, T.W., Cronin, M.T., 2005. The better predictive model: high q2 for the training set or low root mean square error of prediction for the test set? QSAR Comb. Sci. 24 (3), 385–396.
Babaeian, E., Paheding, S., Siddique, N., Devabhaktuni, V.K., Tuller, M., 2022. Short-and mid-term forecasts of actual evapotranspiration with deep learning. J. Hydrol. 612, 128078.
Bergmeir, C., Benítez, J.M., 2012. On the use of cross-validation for time series predictor evaluation. Inf. Sci. 191, 192–213.
Cannas, B., Fanni, A., See, L., Sias, G., 2006. Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. Phys. Chem. Earth Parts A/B/C 31 (18), 1164–1171.
Chen, Z., Zhu, Z., Jiang, H., Sun, S., 2020. Estimating daily reference evapotranspiration based on limited meteorological data using deep learning and classical machine learning methods. J. Hydrol. 591, 125286.
Chen, Z., Wang, G., Wei, X., Liu, Y., Duan, Z., Hu, Y., Jiang, H., 2024. Basin-scale daily drought prediction using convolutional neural networks in Fenhe River basin, China. Atmosphere 15 (2), 155.
Chia, M.Y., Huang, Y.F., Koo, C.H., Ng, J.L., Ahmed, A.N., El-Shafie, A., 2022. Long-term forecasting of monthly mean reference evapotranspiration using deep neural network: a comparison of training strategies and approaches. Appl. Soft Comput. 126, 109221.
da Silva, D.G., de Moura Meneses, A.A., 2023. Comparing long short-term memory (LSTM) and bidirectional LSTM deep neural networks for power consumption prediction. Energy Rep. 10, 3315–3334.
Dong, S., Wang, P., Abbas, K., 2021. A survey on deep learning and its applications. Comput. Sci. Rev. 40, 100379.
Elsner, J., Tsonis, A., 1992. Nonlinear prediction, chaos, and noise. Bull. Am. Meteorol. Soc. 73 (1), 49–60.
Fan, J., Yue, W., Wu, L., Zhang, F., Cai, H., Wang, X., Lu, X., Xiang, Y., 2018. Evaluation of SVM, ELM and four tree-based ensemble models for predicting daily reference evapotranspiration using limited meteorological data in different climates of China. Agric. For. Meteorol. 263, 225–241.
Ferreira, L.B., da Cunha, F.F., 2020. Multi-step ahead forecasting of daily reference evapotranspiration using deep learning. Comput. Electron. Agric. 178, 105728.
Ferreira, L.B., da Cunha, F.F., de Oliveira, R.A., Fernandes Filho, E.I., 2019. Estimation of reference evapotranspiration in Brazil with limited meteorological data using ANN and SVM – a new approach. J. Hydrol. 572, 556–570.
Fijani, E., Barzegar, R., Deo, R., Tziritis, E., Skordas, K., 2019. Design and implementation of a hybrid model based on two-layer decomposition method coupled with extreme learning machines to support real-time environmental monitoring of water quality parameters. Sci. Total Environ. 648, 839–853.
Fisher, J.B., Melton, F., Middleton, E., Hain, C., Anderson, M., Allen, R., McCabe, M.F., Hook, S., Baldocchi, D., Townsend, P.A., 2017. The future of evapotranspiration: global requirements for ecosystem functioning, carbon and climate feedbacks, agricultural management, and water resources. Water Resour. Res. 53 (4), 2618–2626.
Gelete, G., Yaseen, Z.M., 2024. Hybridization of deep learning, nonlinear system identification and ensemble tree intelligence algorithms for pan evaporation estimation. J. Hydrol. 131704.
Ghimire, S., Deo, R.C., Casillas-Pérez, D., Salcedo-Sanz, S., Sharma, E., Ali, M., 2022. Deep learning CNN-LSTM-MLP hybrid fusion model for feature optimizations and daily solar radiation prediction. Measurement 202, 111759.
Goyal, S., Parashar, A., 2018. Machine learning application to improve COCOMO model using neural networks. Int. J. Inform. Technol. Comput. Sci. (IJITCS) 3, 35–51.
Granata, F., Di Nunno, F., 2021. Forecasting evapotranspiration in different climates using ensembles of recurrent neural networks. Agric. Water Manag. 255, 107040.
Huang, Y., 2009. Advances in artificial neural networks–methodological development and application. Algorithms 2 (3), 973–1007.
Jung, D.-H., Lee, T.S., Kim, K., Park, S.H., 2022. A deep learning model to predict evapotranspiration and relative humidity for moisture control in tomato greenhouses. Agronomy 12 (9), 2169.
Karbasi, M., Jamei, M., Ali, M., Malik, A., Yaseen, Z.M., 2022. Forecasting weekly reference evapotranspiration using auto encoder decoder bidirectional LSTM model hybridized with a Boruta-CatBoost input optimizer. Comput. Electron. Agric. 198, 107121.
Kirchgässner, G., Wolters, J., Hassler, U., 2012. Introduction to Modern Time Series Analysis. Springer Science & Business Media.
Lary, D.J., Alavi, A.H., Gandomi, A.H., Walker, A.L., 2016. Machine learning in geosciences and remote sensing. Geosci. Front. 7 (1), 3–10.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. nature 521 (7553), 436–444.
Liu, Y., Yan, X., Du, W., Zhang, T., Bai, X., Nan, R., 2024. Downscaling daily reference evapotranspiration using a super-resolution convolutional transposed network. Water 16 (2), 335.

Morillas, L., Leuning, R., Villagarcía, L., García, M., Serrano-Ortiz, P., Domingo, F., 2013. Improving evapotranspiration estimates in Mediterranean drylands: the role of soil evaporation. Water Resour. Res. 49 (10), 6572–6586.
Narayanan, A., Menneer, T., 2000. Quantum artificial neural network architectures and components. Inf. Sci. 128 (3–4), 231–255.
Oki, T., Kanae, S., 2006. Global hydrological cycles and world water resources. science 313 (5790), 1068–1072.
Quilty, J., Adamowski, J., 2018. Addressing the incorrect usage of wavelet-based hydrological and water resources forecasting models for real-world applications with best practices and a new forecasting framework. J. Hydrol. 563, 336–353.
Richetti, J., Diakogianis, F.I., Bender, A., Colaço, A.F., Lawes, R.A., 2023. A methods guideline for deep learning for tabular data in agriculture with a case study to forecast cereal yield. Comput. Electron. Agric. 205, 107642.
Roy, D.K., 2021. Long short-term memory networks to predict one-step ahead reference evapotranspiration in a subtropical climatic zone. Environ. Process. 8, 911–941.
Saggi, M.K., Jain, S., 2019. Reference evapotranspiration estimation and modeling of the Punjab northern India using deep learning. Comput. Electron. Agric. 156, 387–398.
Schultz, A., Wieland, R., 1997. The use of neural networks in agroecological modelling. Comput. Electron. Agric. 18 (2–3), 73–90.
Schultz, A., Wieland, R., Lutze, G., 2000. Neural networks in agroecological modelling—stylish application or helpful tool? Comput. Electron. Agric. 29 (1–2), 73–97.
Schuurmann, G., Ebert, R.-U., Chen, J., Wang, B., Kuhne, R., 2008. External validation and prediction employing the predictive squared correlation coefficient test set activity mean vs training set activity mean. J. Chem. Inf. Model. 48 (11), 2140–2145.
Sharafi, M., Abdi, E., Mohebbiyan, M., Samadianfard, S., 2024. Prediction of daily evapotranspiration using the strategy of combining tree models with empirical Hargreaves method. Water Soil Sci. 34 (2), 107–119.
Sherstinsky, Alex, 2020. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. Physica D 404, 132306.
Talib, A., Desai, A.R., Huang, J., Griffis, T.J., Reed, D.E., Chen, J., 2021. Evaluation of prediction and forecasting models for evapotranspiration of agricultural lands in the Midwest US. J. Hydrol. 600, 126579.
Vaz, P.J., Schütz, G., Guerrero, C., Cardoso, P.J., 2022. Hybrid neural network based models for evapotranspiration prediction over limited weather parameters. IEEE Access 11, 963–976.
Wang, X.-D., Xu, H.-J., Pan, Y.-X., Yang, X.-M., 2024. Forecasting ecological water demand of an arid oasis under a drying climate scenario based on deep learning methods. Eco. Inform. 102721.
Ye, H., Gao, F., Yin, Y., Guo, D., Zhao, P., Lu, Y., Wang, X., Bai, J., Cao, K., Song, Q., 2019. Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. Eur. Radiol. 29, 6191–6201.
Zhang, L., Zhao, X., Zhu, G., He, J., Chen, J., Chen, Z., Traore, S., Liu, J., Singh, V.P., 2023. Short-term daily reference evapotranspiration forecasting using temperature-based deep learning models in different climate zones in China. Agric. Water Manag. 289, 108498.