



OPEN

## Modelling monthly pan evaporation utilising Random Forest and deep learning algorithms

Mustafa Abed<sup>1</sup>, Monzur Alam Imteaz<sup>1</sup>, Ali Najah Ahmed<sup>2</sup> & Yuk Feng Huang<sup>3</sup>✉

Evaporation is the primary aspect causing water loss in the hydrological cycle; therefore, water loss must be precisely measured. Evaporation is an intricate nonlinear process occurring as a result of several climatic aspects. The purpose of this research is to assess the feasibility of using Random Forest (RF) and two deep learning techniques, namely convolutional neural network (CNN), and deep neural network (DNN) to accurately estimate monthly pan evaporation rates. Month-based weather data gathered from four Malaysian weather stations during the 2000–2019 timeframe was used to train and evaluate the models. Several input attributes (predictor variables) were investigated to select the most suitable variables for machine learning models. Every approach was tested with several models, each with a different set of model aspects and input parameter combinations. The formulated ML approaches were benchmarked against two commonly used empirical methods: Stephens & Stewart and Thornthwaite. Model outcomes were assessed using standard statistical measures to determine their effectiveness in predicting evaporation. The results indicated that the three ML models developed in the study performed better than empirical models and could significantly improve the precision of monthly  $E_p$  estimates even with the identical input sets. The performance assessment metrics also show that the formulated CNN approach was acceptable for modelling monthly water loss due to evaporation with a higher degree of accuracy than other ML frameworks explored in this study. In addition, the CNN framework outperformed other AI techniques evaluated for the same areas using identical data inputs. The investigation's findings in relation to the various performance criteria show that the proposed CNN model is capable of capturing the highly non-linearity of evaporation and could be regarded as an effective tool to predict evaporation.

### Abbreviations

AI	Artificial intelligence
ANN	Artificial neural networks
CNN	Convolutional neural network
CV	Cross-validation
DL	Deep learning
DNN	Deep neural network
$E_p$	Pan evaporation
MAE	Mean absolute error
ML	Machine learning
MSE	Mean square error
NSE	Nash–Sutcliffe efficiency
PCC	Pearson correlation coefficient
$R^2$	Coefficient of determination
RAE	Relative absolute error
RF	Random Forest

<sup>1</sup>Department of Civil and Construction Engineering, Swinburne University of Technology, Hawthorn, Melbourne, VIC 3122, Australia. <sup>2</sup>Department of Civil Engineering, College of Engineering, Universiti Tenaga Nasional (UNITEN), 43000 Selangor, Malaysia. <sup>3</sup>Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Selangor, Malaysia. ✉email: huangyf@utar.edu.my

RH	Relative humidity
RMSE	Root mean square error
$R_s$	Solar radiation
SD	Standard deviation
SDA	Standard deviation actual
SDP	Standard deviation predicted
$S_w$	Wind speed
$T_a$	Mean air temperature
$T_{max}$	Maximum air temperature
$T_{min}$	Minimum air temperature

Evaporation is among the vital aspects that have a pivotal role in regulating the hydrological cycle; forecasting evaporation loss is critically important for water management, irrigation planning, and agricultural models<sup>1–4</sup>. Increased evaporation rate is a significant global warming indicator<sup>5</sup>. Therefore, recording evaporation patterns is critical for monitoring and handling water resources<sup>6</sup>. Evaporation causes significant water loss, impacting water levels in lakes and reservoirs, affecting the water budget. Therefore, before designing irrigation systems and implementing water resource strategies, evaporation losses must be estimated<sup>7</sup>. Reliable evaporation forecasting is critical for hydrological and water resources, enhancement of water use, and water balance. Vapour pressure difference and heat availability affect evaporation rates; these determining factors are affected by meteorological aspects, such as solar radiation, humidity levels, wind speed, air temperature, and air pressure<sup>8–10</sup>. Such factors are also deeply associated with other characteristics like geographical location, seasonal influence, climate type, and time of day. Hence, evaporation is a complex phenomenon with extremely non-linear characteristics.

Evaporation estimation is conducted using indirect and direct techniques, including energy balance, water balance, mass transfer, Penman method, and evaporation pan<sup>11</sup>. The evaporation pan is an extensively used apparatus because it is inexpensive and easy to use<sup>12</sup>. Nevertheless, this is an energy-intensive process affected by numerous meteorological aspects like wind speed and vapour pressure. Moreover, pan evaporimeters cannot be deployed at every required location, specifically those where instruments cannot be installed or managed<sup>13</sup>. Indirect techniques comprise evaporation determination using meteorological information and physical concepts like volume and energy conservation that require precise adjustment based on climate. Accurately determining such meteorological variables is challenging and requires advanced instruments and skilled labour<sup>14</sup>. However, it is known that such techniques cannot offer reliable evaporation data because of intrinsic complications and the non-linear nature of the evaporation process. Considering the inadequate performance levels, such techniques have prompted scientists to develop alternative methods for determining evaporation levels<sup>15</sup>.

**Literature review.** Recently, AI techniques like ANN, M5 model tree (MT), support vector machines (SVM), adaptive neuro-fuzzy inference system (ANFIS), extreme learning machine (ELM), and gene expression programming (GEP) have been used to handle different water engineering and environmental issues<sup>16–21</sup>. Such AI techniques are simpler, more robust and can model complex non-linear processes without significant problems<sup>13,22,23</sup>. Extensive research has been conducted about using AI to forecast different hydrological parameters<sup>24</sup>. Researchers assert that ANN frameworks provide better forecasts than conventional methods. For example, Castellano-Méndez et al.<sup>25</sup> contrasted the Box & Jenkins approach with ANN; the latter provides better runoff simulation performance in terms of precision.

Concerning evaporation forecasting and considering the challenges of practical and conceptual measurement techniques discussed above, several works have been performed using ML approaches with several optimisation works for forecasting pan evaporation<sup>26,27</sup>. They offered specific distinct machine learning approaches for the problem using different input sets concerning existing climatic attributes like wind speed, temperature, humidity, vapour pressure, solar radiation, and sunshine<sup>28,29</sup>. Keskin and Terzi<sup>30</sup> used ANN and Penman models to develop evaporation models. They employed several meteorological aspects as ANN inputs. These researchers indicated that ANNs were superior to the Penman approach for evaporation forecasting. Kişi<sup>31</sup> formulated evolutionary neural networks to estimate pan evaporation for monthly timescales. The results indicated that the formulated models provided better accuracy than empirical methods. Deo et al.<sup>32</sup> researched monthly water loss due to evaporation; they used three machine learning techniques: Relevance Vector Machine (RVM), Extreme Learning Machine (ELM), and Multivariate Adaptive Regression Spline (MARS). Meteorological aspects were employed as independent variables, and RVM was found to be the most effective approach among these. Sudheer et al.<sup>22</sup> formulated an ANN approach for modelling daily evaporation and mentioned that ANN frameworks could be effectively employed to forecast evaporation using climate data. Falamarzi et al.<sup>33</sup> evaluated ANN and wavelet ANN use to forecast daily evaporation. They employed temperature and wind speed data as model inputs. The results indicated that the two frameworks estimated evaporation precisely. Wang et al.<sup>34</sup> estimated daily evaporation using multivariate adaptive regression spline (MARS), least-square support vector regression (LSSVR), fuzzy genetic (FG), multiple linear regression (MLR), and M5 model tree (M5Tree) for eight locations near the Dongting Lake basin in China. The outcomes indicated that FG and LSSVR offer better performance and estimate evaporation with high accuracy. Malik et al.<sup>35</sup> estimated monthly Ep in the central Himalayan region in India using radial basis neural network (RBNN), multilayers perceptron neural network (MLPNN), self-organising map neural network (SOMNN), and co-active neuro-fuzzy inference system (CANFIS). The appropriate input set was selected using the Gamma test. The researchers found that the AI-powered technique could be employed for precise evaporation prediction. Tezel and Buyukyildiz<sup>36</sup> studied the applicability of MLP, RBFN, and e-support

vector regression (SVR) using numerous training methods. When scaled conjugate gradient (SCG) learning was used for ANN and SVR approaches, performance was higher than empirical approaches.

Tree-based Machine learning approaches like RF have been used extensively for water and other environmental modelling during the past decade to estimate aspects like groundwater levels, streamflow, solar radiation, soil moisture, evaporation (e.g., pan and potential evapotranspiration), and suspended sediment. Such methods are relatively straightforward but potent approaches for pattern or trend detection<sup>37,38</sup>; moreover, they offer more computationally efficient for relatively large datasets than other machine learning techniques<sup>39</sup>. Francke et al.<sup>40</sup> employed quantile regression forests (QRF) to estimate suspended sediment concentration for four sub-catchment areas in Spain. QRF results were contrasted against RF, generalised linear model, and the traditional sediment rating curve. The researchers found that QRF and RF are extremely flexible techniques that successfully modelled sediment dynamics. Feng et al.<sup>41</sup> employed the RF framework to predict daily evaporation in southwest China and contrasted it against the GRNN approach. The outcomes indicated that GRNN and RF approach provided acceptable results concerning daily evaporation; RF was marginally superior to GRNN. Recently, DL methods have been used in the machine learning domain and demonstrated success for data evaluation for natural processes, attracting attention concerning time series forecasts<sup>42</sup>. Deep learning is a recent development in the ML paradigm that evolved from near-human to super-human performance for several engineering scenarios. In this class, forecasts are impacted by prior system characteristics; therefore, they may be used for regression and classification problems. Evaporation is intrinsically complex, dynamic, and non-linear; thus, the adaptive evaporation estimation framework must process nonlinear properties. Of the many ANN models specified in the literature, DL can process higher-order non-linear characteristics with better performance concerning time series data and its intrinsic properties for extended durations to enhance forecasting performance<sup>43</sup>. Convolutional Neural Network (CNN) has garnered extensive attention in the deep learning technique domain due to its use in several domains like object recognition<sup>44</sup>, time series categorisation<sup>45</sup>, audio signal classification<sup>46</sup>, and robotic visual and haptic data classification<sup>47</sup>, and weather forecasting<sup>48</sup>. In addition, in the noisy time series context, convolutional networks also reduce data noise and identify useful patterns by building hierarchical structures<sup>49</sup>. It must be noted that several academicians have used CNN for numerous time series prediction fields like solar energy forecast, electrical load estimation, and other scenarios.

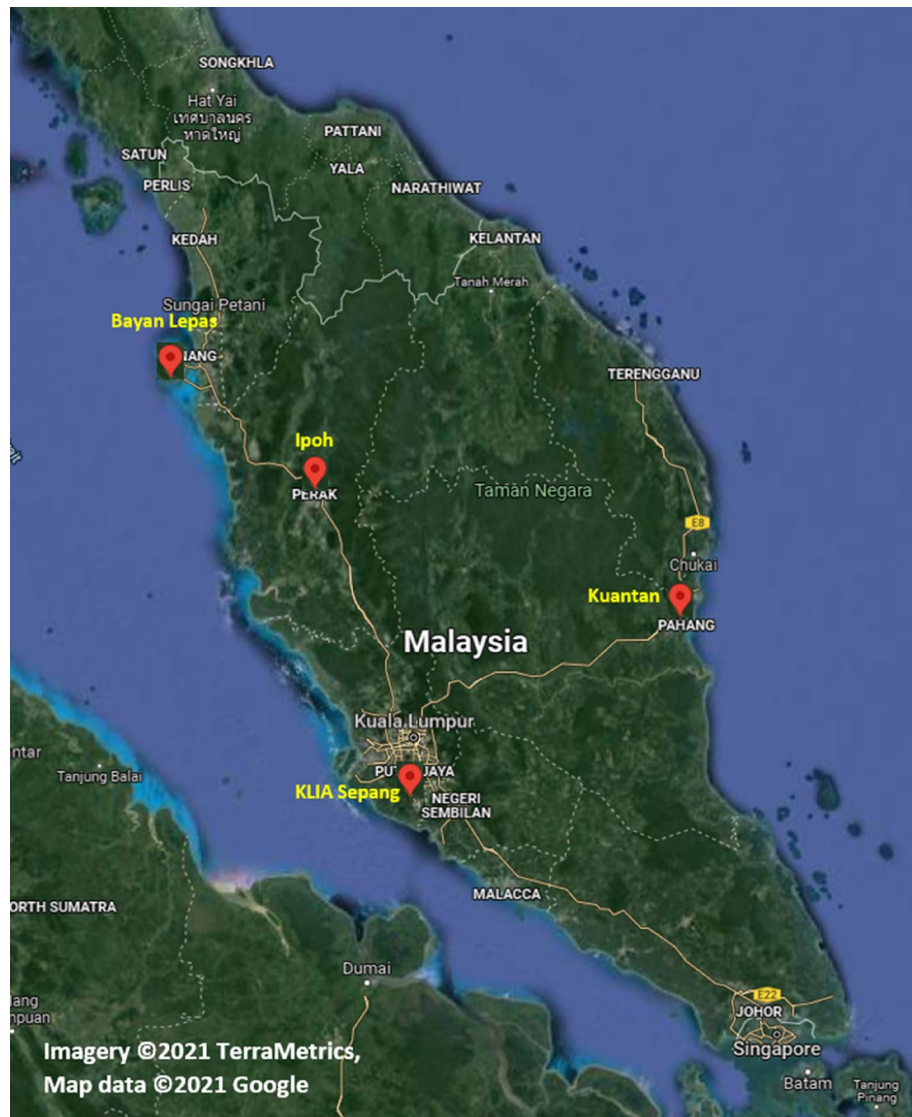
The literature review confirmed that using ANN with appropriate learning methods can suitably model evaporation for numerous locations with superior results than relatively complex conventional approaches<sup>50</sup>. However, identifying and devising efficacious, reliable, and generalised evaporation estimation techniques is still challenging for researchers because of the intricate and non-linear nature of the evaporation process. Among the diverse ANN methods used in the recent past, the cutting-edge DL approach offers immense potential for prediction problems and has outperformed more complex methods. Because prediction is a nonlinear task, the adaptive framework for prediction ought to be nonlinear as well. With the success of DL, the CNN has become extremely advantageous for extracting characteristics from time-series data signals and thus for classification and prediction. The most important aspect of this approach is that it identifies implied recurrent sequences from the series. Moreover, such networks automatically use data to identify features without additional training or prior information. The CNN is powerful in capturing high nonlinear features among the various DL structures reported in the literature. Hence, in the current study, CNN was selected for the monthly pan evaporation forecast.

**Objectives.** This study is intended to assess the predictability and applicability of the CNN model in accurately estimating monthly  $E_p$  rates in four Malaysian regions using weather data for the period 2000–2019. The performance of the CNN model is compared with that of the RF as a powerful tree-based technique and with the DNN model. The models' prediction accuracy is explored under various input combination scenarios. The proposed ML frameworks are contrasted against two widely used empirical methods, namely Thornthwaite and Stephens & Stewart, under identical input combinations. The model's efficiency values are assessed and analysed using standard statistical performance metrics to determine their use in predicting evaporation levels. Furthermore, sufficient analysis would be performed in this study to demonstrate the reliability of the CNN model, with the goal of developing a dependable model for predicting evaporation, which is essential, specifically in water resource management and agricultural planning.

## Study area and data

**Study area.** Malaysia is in the tropical region and receives ample rainfall. Nevertheless, development has spiked water requirements. Additionally, climate change has extended the dry season and increased the evaporation rate from reservoirs. Many consider drought a very intricate but poorly understood natural calamity, impacting people more than other hazards<sup>51</sup>; hence, predicting evaporation is vital. Therefore, this research, which aims to develop accurate models for predicting  $E_p$ , is extremely important, particularly in water resource management and agriculture. The climate monthly data from four meteorological stations situated in Bayan Lepas (longitude 100° 16' E, latitude 5° 18' N, elevation 2.5 m), Ipoh (longitude 101° 06' E, latitude 4° 34' N, elevation 40.1 m), KLIA Sepang (longitude 101° 42' E, latitude 2° 44' N, elevation 16.1 m), and Kuantan (longitude 103° 13' E, latitude 3° 46' N, elevation 15.2 m), managed by the MMD (Malaysian Meteorological Department), are utilised to calibrate and corroborate the recommended predictive models. Figure 1 depicts Malaysia's map where the four stations are situated; Google Maps were used to create this map depicting the studied region.

**Data description.** The propositioned predictive models were built using seven meteorological indicators that include  $T_{max}$ ,  $T_{min}$ ,  $T_a$ , RH,  $S_w$ ,  $R_p$ , and  $E_p$ . The data set consisted 19 years of day-to-day reports from 2000 to 2019. The statistical parameters recorded every month pertaining to the quantified meteorological data for the



**Figure 1.** Location of case study [Imagery ©2021 TerraMetrics, Map data ©2021 Google].

four above-mentioned stations are listed in Table 1. Moreover, Fig. 2 illustrates the monthly variation of every weather parameter for the duration 2000 to 2019.

In the table, the  $X_{min}$ ,  $X_{max}$ ,  $X_{mean}$ ,  $S_x$ ,  $C_x$ , and  $C_v$  represent the minimum, maximum, mean, standard deviation, skewness, and coefficient of variation of the weather parameters, respectively. It is apparent from this table data that the Ep minimum value was measured at Kuantan station, whereas the maximum value was recorded in the Bayan Lepas station. This might be due to the rate of relative humidity, which is inversely associated to evaporation; Kuantan station has the highest relative humidity rate, and Bayan Lepas station represents the lowest rate. Conversely, the coefficient of variation and maximum skewness of Ep were also measured in the Bayan Lepas station, while the minimum value was recorded in Ipoh. A positive value of skewness implies that the information is not symmetric and does not adhere to the normal distribution.

**Partitioning of data and input selection.** Selecting the suitable predictors is one of the most crucial steps in developing a robust predictive model<sup>52</sup>; different input combinations of meteorological parameters were examined in this study to successfully plot input–output model and improve the predictive ability of ML models. This will enable a better practical comprehension of how every input parameter affects the evaporation estimate in that region<sup>53</sup>. There are certain conscious choices for choosing these combinations. First, for the purpose of comparison, input variables to the models of machine learning (RF, DNN, and CNN) were chosen according to the required meteorological aspects in the two proposed empirical models (Thorntwaite and Stephens & Stewart). Second, the input variables (predictors) were chosen with reference to the PCC<sup>54</sup>. The Pearson correlation method is the test statistics that quantifies the statistical correlation, or association, among two continuous parameters. It is identified as the best technique of measuring the correlation between parameters of interest since it is based on the covariance method<sup>55</sup>. It gives data about the association or correlation magnitude and the

Station	Dataset	Unit	$X_{mean}$	$S_x$	$C_v$	$C_x$	$X_{min}$	$X_{max}$
Bayan Lepas	$T_{max}$	°C	31.89	0.75	2.36	0.52	30.15	34.78
	$T_{min}$	°C	24.94	0.51	2.07	0.47	23.73	26.90
	RH	%	78.74	4.16	5.29	-1.05	64.66	85.92
	$S_w$	m/s	1.93	0.34	17.99	0.86	1.20	3.25
	$R_s$	MJ m <sup>-2</sup>	18.32	2.14	11.69	1.06	12.72	28.49
	$E_p$	mm	3.89	0.66	17.11	1.17	2.70	6.27
Ipoh	$T_{max}$	°C	32.99	0.84	2.55	0.35	30.59	35.80
	$T_{min}$	°C	23.93	0.53	2.24	-0.03	22.40	25.40
	RH	%	80.54	3.92	4.87	-0.53	68.33	88.71
	$S_w$	m/s	1.52	0.32	21.18	-1.28	0.62	2.15
	$R_s$	MJ m <sup>-2</sup>	17.79	1.49	8.40	-0.19	13.31	21.72
	$E_p$	mm	4.29	0.49	11.47	0.007	3.21	5.69
KLIA Sepang	$T_{max}$	°C	32.20	0.81	2.54	0.63	30.29	34.77
	$T_{min}$	°C	24.42	0.49	2.02	0.10	23.24	25.69
	RH	%	79.62	4.13	5.19	-0.82	63.51	87.51
	$S_w$	m/s	1.87	0.27	14.62	0.45	1.15	2.81
	$R_s$	MJ m <sup>-2</sup>	17.55	2.38	13.59	0.59	11.12	24.76
	$E_p$	mm	4.17	0.48	11.66	1.13	3.20	6.12
Kuantan	$T_{max}$	°C	32.17	1.24	3.88	-0.63	28.52	34.89
	$T_{min}$	°C	23.71	0.64	2.70	-0.64	21.14	25.53
	RH	%	84.29	3.01	3.58	0.33	77.33	92.39
	$S_w$	m/s	1.64	0.30	18.49	0.64	0.91	2.65
	$R_s$	MJ m <sup>-2</sup>	17.26	2.16	12.56	-0.47	11.73	22.39
	$E_p$	mm	3.79	0.53	13.64	-0.26	2.69	5.13

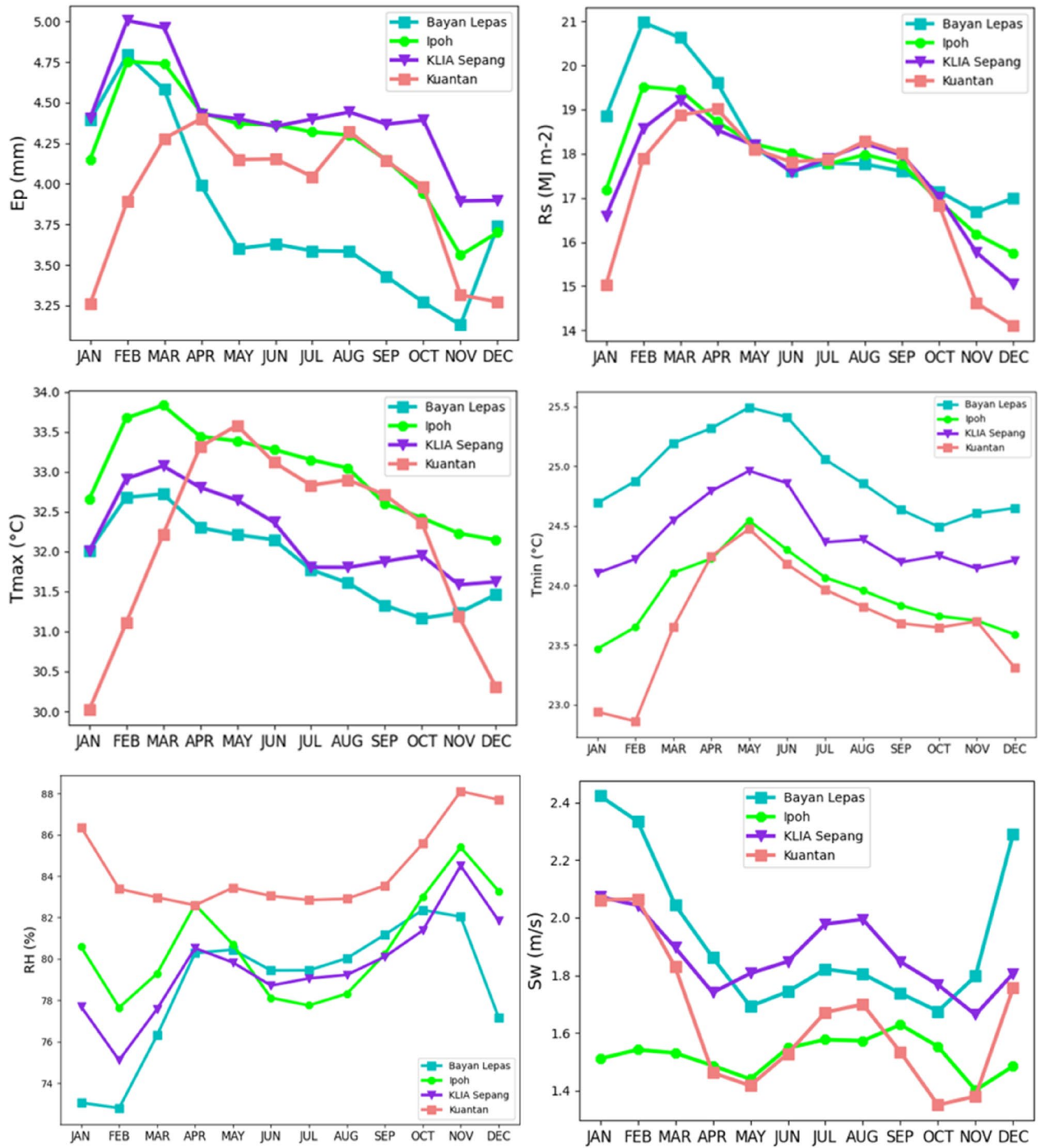
**Table 1.** Various meteorological variables and their descriptive statistics.

direction of the correlation. The two parameters can be negatively or positively associated and there is no relationship among the two parameters if the PCC is 0. To show the applicable features of the environmental parameters to estimate monthly evaporation, the PCC interpretations and ranges are used as displayed in Table 2. The PCC were employed to find the meteorological parameters showing the greatest effect on the estimates of evaporation, and the results are shown in Table 3.

The outcomes are listed in Table 3, indicating that the  $T_{max}$ ,  $T_{min}$ , RH,  $S_w$ ,  $R_s$  were all related to a certain extent with  $E_p$  and therefore can play a crucial role in predicting the evaporation parameter for the data gathered at all stations. Particularly, at all stations, the  $T_{max}$  and RH parameters have the strongest relationship with  $E_p$ . Thus, the  $T_{max}$  and RH will be employed in all input combinations in order to increase the  $E_p$  estimation accuracy. Earlier studies also suggested that  $T_{max}$ ,  $T_{min}$ , RH,  $S_w$ , and  $R_s$  are some of the most significant predictors of evaporation<sup>56,57</sup>.

The current research has also evaluated the effects of the input parameter  $E_p$  in improving the prediction accuracy for evaporation. In this regard, the records of input data were chosen with reference to how the previous records were associated with the estimated output value. As illustrated in Fig. 3, at each of the stations, the autocorrelation examination for the recorded time series on monthly basis for the  $E_p$  rate showed that the correlation declined significantly once it went beyond the previous second lag-time record. This shows that the previous record of second evaporation rate affected the evaporation rate at any time. Therefore, based on the past pan evaporation rate records with the advantage of the correlation analysis, the highest lag times of two previous records were employed as the model input when building the proposed models on monthly basis.

Accordingly, in the current study, nine different input scenarios were considered for the models (Table 4). Each climatic data set was divided into two sets, in which 80% was employed for model calibration (training) while 20% was used for validation (testing). Thus, the dataset was partitioned by taking the initial years for training and the remaining years for testing. However, the evaluation of ML approaches is extremely sensitive to the adopted data partitioning scheme. Therefore, the k-fold CV technique would be used. Despite the high computational cost associated in the CV method, it is regarded as one of the reliable prevention methods against overfitting<sup>58</sup>. The current study intends to perform a comprehensive assessment for testing AI ability and using practical models for predicting  $E_p$  levels on a monthly basis in the Bayan Lepas, Ipoh, KLIA Sepang, and Kuantan regions.



**Figure 2.** Monthly variations of Ep and related meteorological parameters used in this study.

PCC ranges	Analysis
0.00 < 0.09	Insignificant
0.10 < 0.19	Weak
0.20 < 0.39	Moderate
0.40 < 0.59	Moderately strong
0.60 < 0.79	Strong
0.80 < 1.00	Extremely strong

**Table 2.** Ranges and analysis of the Pearson correlation coefficient (PCC).

	Tmax	Tmin	RH	Sw	Rs	Ep
Tmax	1					
Tmin	0.713293	1				
RH	-0.67121	-0.2172	1			
Sw	0.383116	0.103138	-0.75333	1		
Rs	0.563469	0.196108	-0.54899	0.358648	1	
Ep	0.714932	0.413693	-0.78168	0.682437	0.704246	1

	Tmax	Tmin	RH	Sw	Rs	Ep
Tmax	1					
Tmin	0.598635	1				
RH	-0.64359	-0.29279	1			
Sw	0.027185	-0.15832	-0.40896	1		
Rs	0.74009	0.291002	-0.60416	0.31927	1	
Ep	0.762407	0.499372	-0.88127	-0.51225	0.643995	1

	Tmax	Tmin	RH	Sw	Rs	Ep
Tmax	1					
Tmin	0.655777	1				
RH	-0.62452	-0.36929	1			
Sw	0.278195	0.15671	-0.44816	1		
Rs	0.500622	0.402229	-0.4582	0.391104	1	
Ep	0.688703	0.410466	-0.70967	0.502534	0.665399	1

	Tmax	Tmin	RH	Sw	Rs	Ep
Tmax	1					
Tmin	0.707286	1				
RH	-0.56873	-0.1759	1			
Sw	-0.43197	-0.49138	-0.14552	1		
Rs	0.701106	0.320207	-0.6274	0.041572	1	
Ep	0.702119	0.415635	-0.79619	-0.68268	0.67207	1

**Table 3.** Pearson correlation coefficient values between the meteorological variables measured at Bayan Lepas, Ipoh, KLIA Sepang, and Kuantan stations.

## Methodology

**Empirical models used for monthly Ep prediction.** In this research, Stephens & Stewart and Thornthwaite were selected for comparison as the two empirical techniques, as they are regarded to be widely employed methods<sup>59</sup>, taking into account the number of meteorological inputs required as well as the availability of the data.

*Stephens & Stewart.* This technique is also commonly referred as the ‘Fractional Evaporation-Equivalent of Solar Energy’ approach by Stephens & Stewart<sup>60</sup>. As presented in Eq. (1), Stephens & Stewart suggested that by employing measured radiation  $Q_s$ , better results were achieved when there is availability of data and it also allows correlating with temperature:

$$Ep = (0.0082Ta - 0.19) \left( \frac{Q_s}{1500} \right) \times 25.4, \quad (1)$$

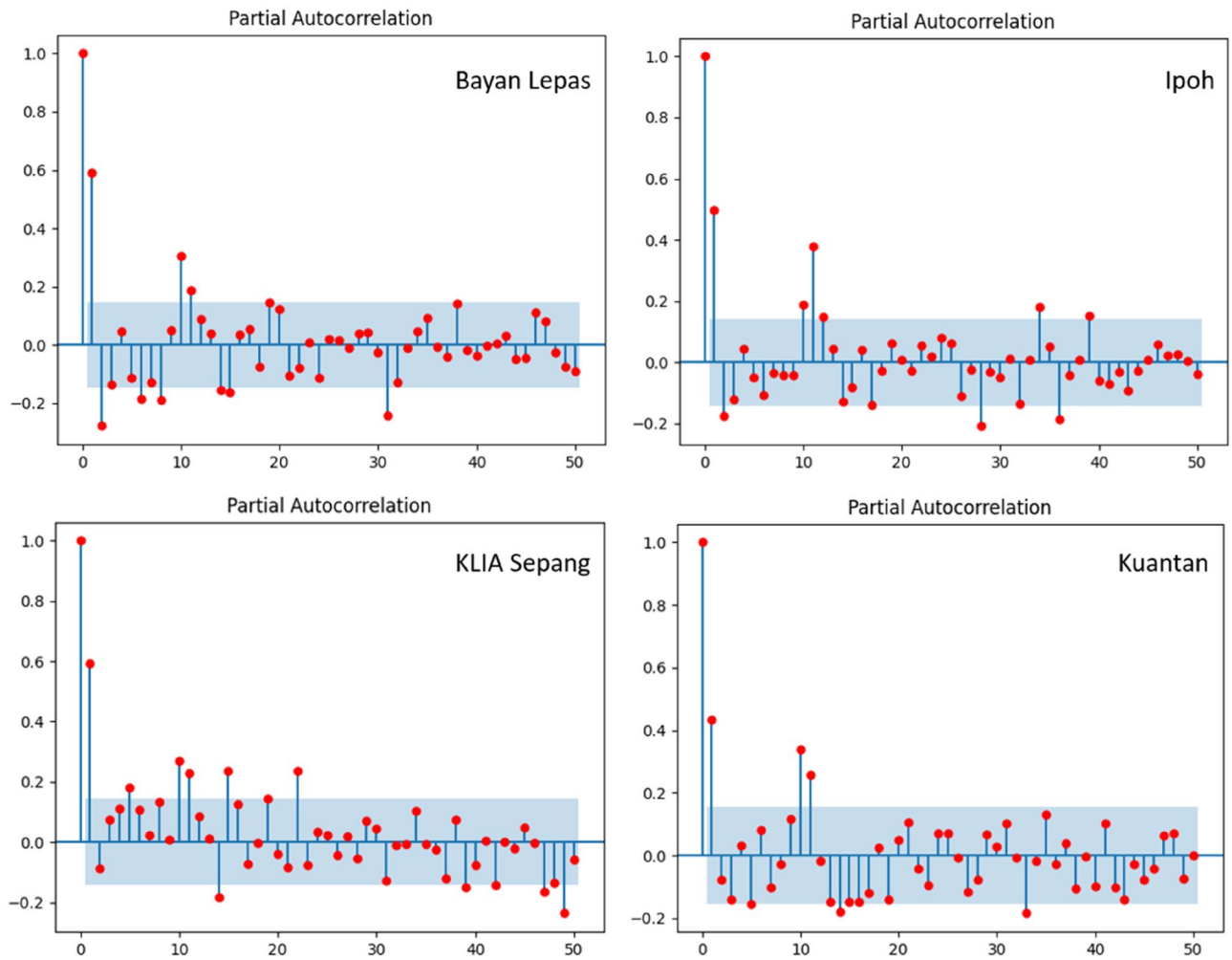
where  $Ta$ ,  $Ep$ , and  $Q_s$  represent mean air temperature (Fahrenheit), evaporation (mm), and solar radiation ( $\text{cal cm}^{-2} \text{day}^{-1}$ ). Stephens & Stewart also recommended carrying out additional research in other regions to set such relationships under different weather conditions.

*Thornthwaite.* Thornthwaite<sup>61</sup> employed practical data to identify the relationship that exists between mean monthly temperature ( $Ta$ ) and probable evaporation ( $Ep$ ), and then set standardisation to a 30-day month with 12 h of sunlight each day. The potential evaporation ( $Ep$ ) is calculated by employing Thornthwaite technique; the following expression is employed to calculate the Monthly Thornthwaite Heat Index ( $I$ ):

$$i = \left( \frac{Ta}{5} \right)^{1.514}, \quad (2)$$

where  $Ta$  represent mean monthly temperature ( $^{\circ}\text{C}$ ).

The Annual heat index ( $I$ ) is calculated as the sum of the Monthly Heat Indices ( $i$ ):



**Figure 3.** Partial autocorrelation for Bayan Lepas, Ipoh, KLIA Sepang and Kuantan stations (Monthly).

No.	Model			Input combinations
	RF	DNN	CNN	
1	RF-1	DNN-1	CNN-1	$T_a$
2	RF-2	DNN-2	CNN-2	$T_a, R_s$
3	RF-3	DNN-3	CNN-3	RH
4	RF-4	DNN-4	CNN-4	RH, $T_{max}$
5	RF-5	DNN-5	CNN-5	RH, $T_{max}$ , $T_{min}$
6	RF-6	DNN-6	CNN-6	RH, $T_{max}$ , $T_{min}$ , $S_w$
7	RF-7	DNN-7	CNN-7	RH, $T_{max}$ , $T_{min}$ , $R_s$
8	RF-8	DNN-8	CNN-8	RH, $T_{max}$ , $T_{min}$ , $R_s$ , $S_w$
9	RF-9	DNN-9	CNN-9	RH, $T_{max}$ , $T_{min}$ , $R_s$ , $S_w$ , Ep

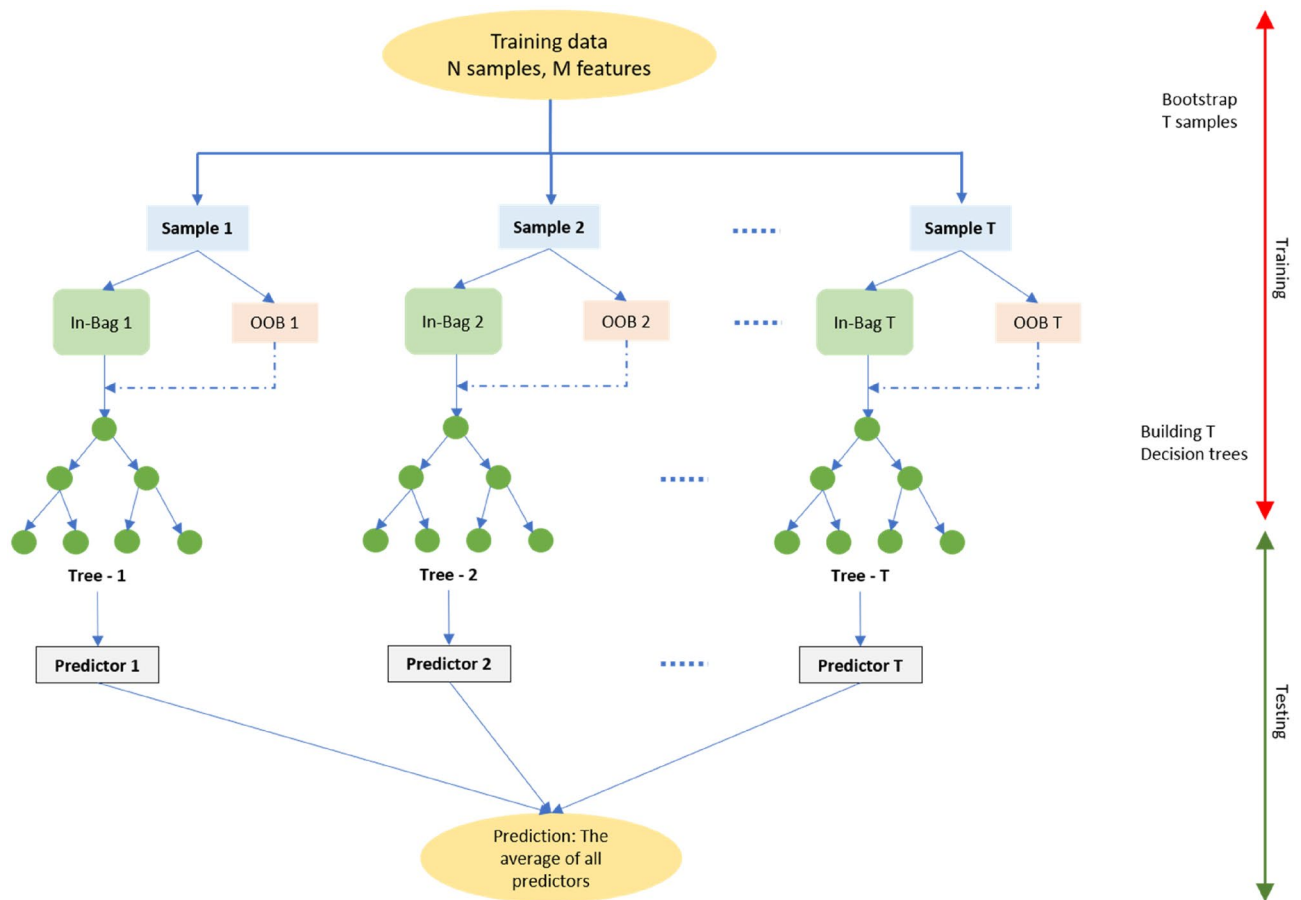
**Table 4.** Input combinations of meteorological variables used for ML models.

$$I = \sum_{i=1}^{12} i. \tag{3}$$

The potential evaporation  $Ep$  for each month is calculated using the following equation:

$$Ep = 16 \cdot \left( \frac{10 \cdot T_a}{I} \right)^a, \tag{4}$$





**Figure 4.** General architecture of the RFs model.

where  $a$  is:

$$a = (675 \times 10^{-9} \times I^3) - (771 \times 10^{-7} \times I^3) - (1792 \times 10^{-5} \times I) + 0.49239. \quad (5)$$

$Ep$  for a given month is given by the expression:

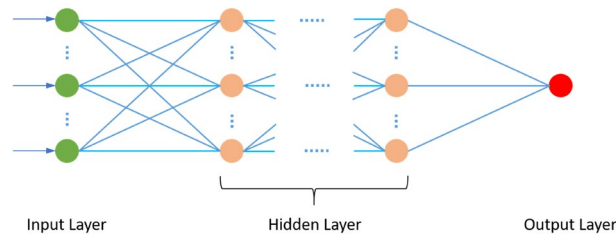
$$Ep = Ep_{Obtained} \cdot \frac{N}{12} \cdot \frac{d}{30} \text{ (mm)}. \quad (6)$$

$N$  and  $d$  denote the number of theoretical monthly sunshine hours and days in the month, respectively.

**ML models used for monthly  $Ep$  prediction.** Three ML frameworks were included in the current study to estimate evaporation, i.e., RF, DNN, and CNN. The TensorFlow framework geared with an NVIDIA GeForce GTX 1080 Ti GPU was employed to conduct training and testing of the machine learning models.

**Random Forest (RF).** The Random Forest algorithm is an effective tree-based ensemble learning algorithm, which is known for its excellent performance. It has a broad range of applications, including regression, classification as well as unsupervised learning<sup>62</sup>. The RFs model was put forward by Breiman<sup>63</sup>, which employed Breiman's 'bagging' idea to ensemble a set of decision trees that possess controlled variation. The data set excluded in the development of the model signified as out-of-bag (OOB) samples is used to assess the general problems (Fig. 4). This also offers a quantitative measurement pertaining to contribution of each input auxiliary data towards the prediction step, referred as RF variable importance<sup>64</sup>. The functioning of Random Forest algorithm in general follows these steps: (i) collect and then re-sample the original training data several times; (ii) select a random set of features for every re-sampling step; (iii) estimate a decision tree based on a re-sample and a random set of features; (iv) to obtain a single decision tree, a set of estimated decision trees is gathered. It can be noted that RF is rather insensitive towards noise as well as overtraining. It has been broadly employed to solve complicated as well as non-linear hydrological engineering issues<sup>65,66</sup>. Additional details about the random forest model theories can be noted in<sup>63</sup>.

In this study, different hyperparameters were employed in RF in order to determine the best ones that can achieve the highest accuracy with regards to prediction, such as:



**Figure 5.** The basic structure of DNN.

1. *The total number of trees needed to generate the forest (Ntree)* This parameter is regarded to be a determinant factor when it comes to conducting predictions with RF.
2. *The tree's maximum depth* With regards to Random Forest, the maximum depth of a tree refers to the longest path between the leaf node and the root node.
3. To identify the best split, the following features need to be kept in mind:
  - max\_features {"auto", "sqrt", "log2"}.
  - If "auto", then max\_features = n\_features.
  - If "sqrt", then max\_features = sqrt (n\_features).
  - If "log2", then max\_features = log2 (n\_features).

*Deep neural network (DNN).* In the deep learning field, DNN are regarded to be a key technique<sup>42</sup>. The fundamental framework has been built by considering the brain's functioning and biological structure to enable machines to achieve intelligence that is more human-like. The basic version pertaining to DNN represents a hierarchical collection of neurons that transmit messages to other neurons as per the input, thus resulting in the development of a complex network learning based on the feedback mechanism. Figure 5 shows the typical structure pertaining to DNN, which includes one input layer, one output layer and numerous hidden layers. As shown in Fig. 5, the balls denote the neurons, wherein each link that exists between neurons is represented by a cause-effect chain that can be trained and learned. The layers remain fully connected, in which any particular neuron in one-layer stays connected to each of the neuron in the next layer. The entire DNN model is made up of a linear function outlined in Eq. (7) as well as an activation function as shown below:

$$a = \sum w_i x_i + b_i, \quad (7)$$

where  $x_i$  represents the input value pertaining to each neuron;  $w_i$  denotes the coefficient pertaining to linear relationship and  $b_i$  defines the bias. Presuming there are  $L$  hidden layers with regards to the DNN, the output value calculation can be represented as follows:

$$f(x) = f[a^{L+1}(h^L(a^L(\dots(h^2(a^2(h^1(a^1(x))))))))]a^L(x) = W^L + b, \quad (8)$$

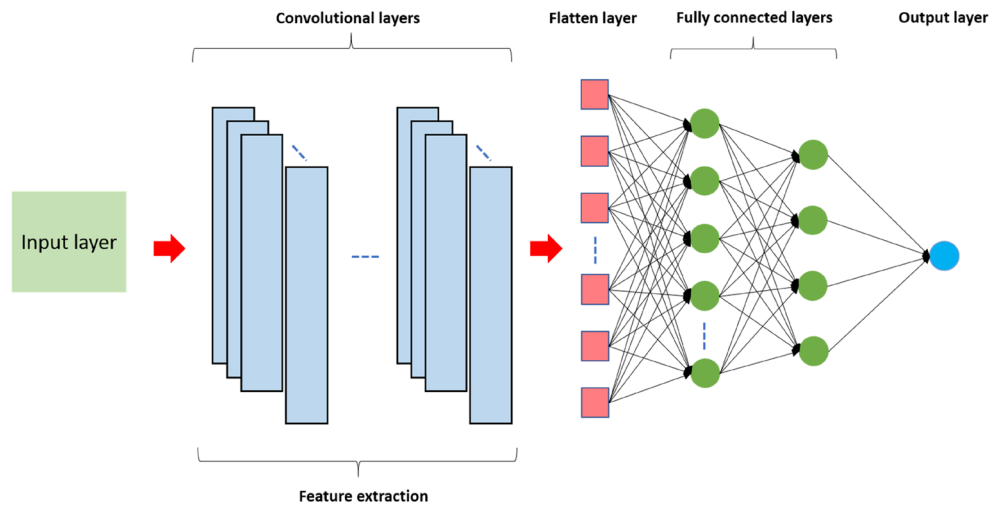
where  $L$  denotes the  $L$ th layer;  $x$  signifies the matrix of input variables;  $b$  and  $W$  indicate high dimensional matrix and  $f(x)$  indicates the introduced activation function to boost the nonlinearity pertaining to the neural network in order to approximate any nonlinear function with regards to numerous nonlinear models. Amongst all of these activation functions, the rectified linear unit (ReLU) activation function, i.e.  $\text{ReLU}(x) = \max(x, 0)$ , has now become the most popular activation functions employed in the deep learning literature as well as applications<sup>67</sup>.

Determination of the values of  $W$  and  $b$  is determined automatically by taking into account the minimum value pertaining to the loss function in the training process. The difference that exists between the actual and predicted values is determined by employing the loss function. The model's robustness gets better when there is a smaller value of loss function. Finally, the output layer is regarded to be the final layer of the network. In this research work, testing of different hyperparameters is done to choose the best architecture that can offer the highest evaluation metrics that will help determine the DNN's optimal structure. The hyperparameters include: (1) The total number of fully connected layers, (2) kinds of activation functions that exist amongst layers, (3) percentage of dropout as well as number of dropout layers, (4) loss function, (5) batch size, (6) optimiser, (7) number of epochs and (8) Learning rate. The put forward DNN model's best architecture with regards to prediction of evaporation includes the following layers:

- (1) Fully connected layers with 64 nodes and ReLU activation function.
- (2) Dropout with 0.1%.
- (3) Fully connected layers with 128 nodes and ReLU activation function.
- (4) Dropout with 0.1%.
- (5) Fully connected layers with 1 node and Linear activation function.

The final hyperparameters are:

1. The learning rate: 0.001.



**Figure 6.** General architecture of the CNN model.

2. Loss Function: Mean Square Error (MSE).
3. Optimizer: ADAM.
4. Epochs: 500.
5. Batch size: 8.

**Convolutional neural network (CNN).** CNN is a renowned and extensively utilised deep learning structure. First recommended by LeCun et al.<sup>68</sup>, CNNs are still a broadly deployed model for image processing and examination due to their capability to mine and decompose features and secure spatial correlations between data in one or two dimensions<sup>69</sup>. Convolutional neural network usually pertains to a 2-dimensional CNN, which is typically utilised for image classification. There are other kinds of CNNs like 1-dimensional (1D-CNN) and 3-dimensional (3D-CNN) which are also utilised in real-life engineering applications. Notably, all CNNs possess the same attributes and follow the same methodology. However, the key dissimilarity is the input data dimensionality and the way the filter (feature detector) moves over the data. In this work, we utilised 1D-CNN for pan evaporation estimation because of its advanced performance and minimal computational intricacy. CNNs comprises two key parts<sup>70</sup>: the first comprises convolutional filtering for mining attributes hierarchically and the second is a fully-connected layer for computing the output value from manifold input values comprising fully-connected neuron layers. The fully connected layers are quite similar to the multilayer perceptron (MLP) layers. The MLP is a feed forward neural network which utilises stochastic gradient descent backpropagation algorithmic for network optimisation. In fact, ordinary artificial neural networks (ANNs) solely comprise the second part; thus, the feature extraction stage is the key difference between CNNs and normal ANNs.

The CNN design generally encompasses an input layer, an output layer and few random numbers of hidden layers among them. A typical CNN setup is depicted in Fig. 6. The input layer is responsible for receiving the signal (input data) as well as transmitting it to the hidden layer(s). Hidden layers can be defined as the computational engine pertaining to the model. These could include one or more dropout layer, convolutional 1D layer, max-pooling layer as well as a flatten layer based on the problem. The CNN's chief building block is the convolutional layer that includes one-dimensional filters/kernels that enable extracting the features via the input signal, an activation function for establishing neurons' threshold limit and kernel size to denote the filter length. There are many commonly utilised activations functions like the ReLU, tanH, Softmax, and Sigmoid. Each of these have a particular use. The hidden information in the input data can be identified and excerpted via convolutional filters. Towards the end of the convolution layers, the learning features are generally flattened to a single long vector array and tend to pass via fully connected layers prior to employing the output layer for prediction. The flatten layer transforms the convolutional/pooling/dropout layers' output to one dimension and then transmits the data to the output layer. To the neurons in the network, the dropout layer (should it be employed) randomly assigns zero weights, making it less sensitive to minor variation, thereby enhancing the model's accuracy regarding unseen data. The 1D-CNN's last layer would be the output layer that contains one neuron for yielding the desired output. To summarise, there exist three kinds of layers which constitute the CNN: the convolutional layers, fully connected (FC) layers, and pooling layers. Once these layers are arranged, a CNN architecture would be created.

In this work, many meteorological variables, such as  $T_{max}$ ,  $T_{min}$ , RH,  $S_w$  and  $R_s$  were applied to CNN to estimate the pan evaporation rate. Iterative parameter tuning helped CNN fit the dataset. To determine the precise CNN structure, several hyperparameters were evaluated to determine the optimal structure to offer the most precise assessment metrics. These hyperparameters comprise convolutional layer count, layer-specific feature map count, filter size, pooling layer category, activation function categories between layers, dropout percentage and numbers, fully-connected layer count, loss function, learning rate, epoch count, batch size, and optimiser. Typically, CNN is built using dense and convolutional layers. Pooling layers might be included in such networks;

the layers are inserted between convolutional layers to decrease problem dimensions and identify critical features. Nevertheless, this study does not consider pooling layers because excess parameter count is tolerable for time series forecasts, and recent studies are critical about the need for pooling layers<sup>71</sup>. Moreover, researchers assert that adequately sized convolutional layers suffice for networking function without adding additional layers<sup>71</sup>. The sequential model is typically used for Python programming, and it was used for this step too. It provides a straightforward technique to create a CNN structure using Keras since it facilitates building the structure based on layers. During CNN training, the objective is to optimise the loss function representing the objective in the neural network structure. The function is based on MSE. This study also employed the dropout technique to reduce overfitting. Dropout is a widely-used regularisation technique (create a more representative CNN weight range by creating a new scale), and the values were 0.1 and 0.2. A batch size of 16 and 500 epochs were chosen for training the model based on the above architectural configuration and several trials. Adam algorithm<sup>72</sup> was employed to adjust network weights to reduce loss function and determine network performance with a learning rate of 0.001 and momentum rate of 0.7.

The one-dimensional CNN structure proposed in this study comprises the following layers for optimal performance concerning evaporation prediction:

- (1) CNN with one convolutional layer and 32 filter with kernel\_size = 2 and activation = 'relu'.
- (2) Dropout with 0.2%.
- (3) Flatten layer (used as a connection between Convolution and the Dense layers).
- (4) Fully connected layers with 128 nodes and ReLU activation function.
- (5) Dropout with 0.1%.
- (6) Fully connected layers with 256 nodes and ReLU activation function.
- (7) Dropout with 0.1%.
- (8) Fully connected layers with 1 node and Linear activation function.

The final hyperparameters are:

1. Learning rate: 0.001.
2. Loss function: MSE.
3. Optimizer: ADAM.
4. Epochs: 500.
5. Batch size: 16.

**Performance evaluation.** Choosing the appropriate performance indicators is crucial since every indicator has its own properties. In addition, knowing the strengths of each statistical measure can provide a better understanding of how the model perform. Therefore, in this study, model predictive performance was evaluated by utilising numerous well-known statistical indicators. These indicators are defined below:

- (1)  $R^2$  the coefficient of determination informs the correlation between the real and estimated outputs; it has a value range of 0–1 (both limits included). Zero indicates a random framework, while one represents optimal fit.  $R^2$  is very popular and makes comparing models easier and more consistent. It attempts to measure how well a regression model is fit a dataset, providing evaluators with an instant understanding of the model's performance.

$$R^2 = \frac{\sum_{i=1}^n (y - \bar{y})(\hat{y} - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y - \bar{y})^2 \sum_{i=1}^n (\hat{y} - \bar{\hat{y}})^2}} \quad (9)$$

- (2)  $MAE$  the absolute difference between the actual and predicted output. High errors caused by outliers are not penalised by MAE. Furthermore, it provides a consistent indicator of how precise the model performs.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y - \hat{y}| \quad (10)$$

- (3)  $MSE$  the average squared difference between predicted and actual output. By squaring the errors, the MSE penalises the model for having large errors. Furthermore, for minor errors, it efficiently converges to the minima.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y - \hat{y})^2 \quad (11)$$

- (4)  $RMSE$  it is the square root of the average value of error squares concerning the real and estimated values. In assessing the performance of a regression model, RMSE is more commonly used than MSE. In addition, RMSE is straightforward and easily distinguishable. RMSE has the added benefit of penalising large errors, making it more acceptable.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y - \hat{y})^2}{n}} \quad (12)$$

- (5) *RAE* the difference between real and forecasted values are gathered and normalised. *RAE* is reliable in some cases because it protects against outliers.

$$RAE = \frac{\sum_{i=1}^n |y - \hat{y}|}{\sum_{i=1}^n |y - \bar{y}|} \quad (13)$$

- (6) *NSE* it represents a normalised metric determining the relative intensively of residual variance (noise) when determined against the calculated variance (information). The *NSE* is still widely used in hydrologic modelling, in part since it normalises performance of the model into an understandable scale.

$$NSE = 1 - \frac{\sum_{i=1}^n (y - \hat{y})^2}{\sum_{i=1}^n (y - \bar{y})^2} \quad (14)$$

where  $n$  is sample count,  $y$  denotes the true output,  $\hat{y}$  denotes the predicted values, and  $\bar{y}$  is the true output average.

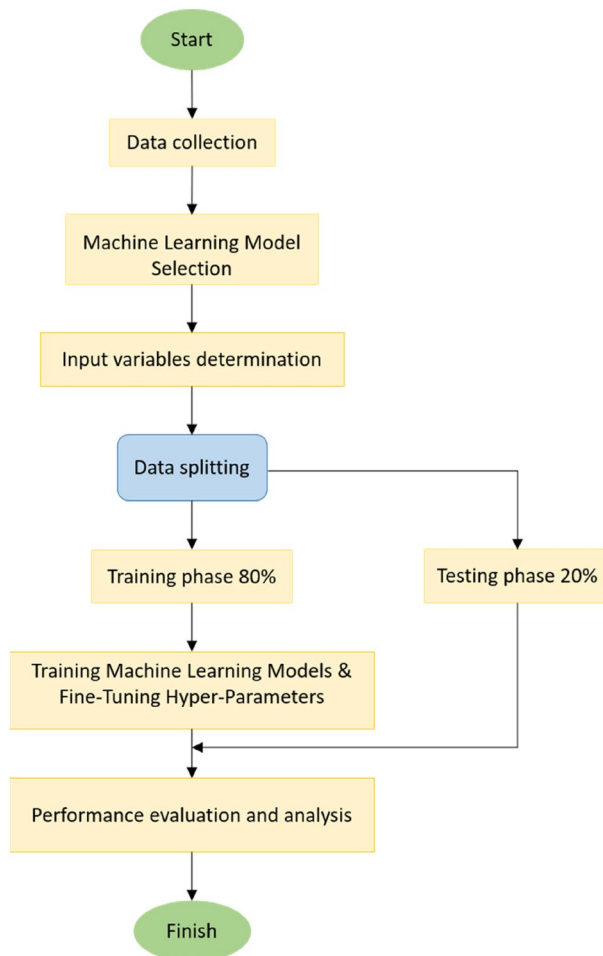
- (7) *Taylor diagram (TD)* Besides the above-mentioned statistical factors, Taylor diagram<sup>73</sup> was also used to calculate the accuracy of the modelling methods taken into consideration and their extent of similarity. The diagram is normally used in climate-based studies<sup>74</sup>. These diagrams can underline the accuracy of models' estimates by comparing the predicted and measured values by visualising a series of elements on a polar plot. The diagram's azimuth angle illustrates the correlation coefficient between the predicted and measured values, whereas the standard deviation value of the modelled data from observations is shown by the radial distance from the origin.

As a conclusion to the performance and training evaluation procedures for the ML models that are proposed, a flow chart is devised which is displayed in Fig. 7. The detailed procedure employed in this approach has been illustrated in the flow chart.

## Results and discussion

**Estimation of monthly Ep using empirical models.** As previously mentioned, monthly Ep was estimated using two empirical models, which include radiation-based and temperature-based models. The values relating to  $R^2$ , MSE, MAE, NSE, RAE and RMSE are recorded in Table 5, with respect to the two models used to estimate Ep in Bayan Lepas, Ipoh, KLIA Sepang and Kuantan stations. As indicated by the statistical values shown in Table 5, greater prediction accuracy was noticed with the model based on radiation (Stephens & Stewart) in comparison with the temperature-based model. Above all, the highest  $R^2$  values (0.620, 0.649, 0.580, and 0.696) and the minimum RMSE values (0.409, 0.292, 0.314, and 0.292) were observed in Stephens & Stewart model for all stations. However, in the Thornthwaite model, values of RMSE increased by approximately average 16%, and the corresponding  $R^2$  reduced by approximately average 33%. The performance values listed in Table 5 clearly suggest that the Stephens & Stewart model surpassed the Thornthwaite model. It could be due to the inclusion of solar radiation, which generally includes an improvement over only the temperature-based estimation<sup>53</sup>. In Figs. 8, 9, 10 and 11, projected values related to monthly Ep with respect to both the empirical models are plotted against the values measured at stations Bayan Lepas, Ipoh, KLIA Sepang and Kuantan, respectively.

**Estimation of monthly Ep using ML models.** Table 6 displays the statistical outcomes related to three ML models with the aim to estimate monthly Ep using nine input combinations with respect to meteorological parameters for Bayan Lepas, Ipoh, KLIA Sepang and Kuantan stations. For every ML model, the optimum statistical parameters have been shown in bold. As can be seen in Table 6, there is a noteworthy difference between the estimation accuracy of monthly Ep based on model type and input combination. According to the statistical values, for different input combinations, with respect to the three machine learning models, the CNN-9 model ( $R^2=0.970$ , MAE=0.071, MSE=0.008, RMSE=0.092, RAE=0.138, NSE=0.980) at the Bayan Lepas station, ( $R^2=0.980$ , MAE=0.053, MSE=0.004, RMSE=0.069, RAE=0.132, NSE=0.981) at the Ipoh station, ( $R^2=0.965$ , MAE=0.079, MSE=0.008, RMSE=0.091, RAE=0.214, NSE=0.966) at the KLIA Sepang station, and ( $R^2=0.962$ , MAE=0.084, MSE=0.010, RMSE=0.103, RAE=0.198, NSE=0.962) at the Kuantan station offered better performance than the DNN and RF models. In addition, as previously stated, the k-fold CV technique has been used. Cross-validation is a reliable method for preventing overfitting. The primary configuration variable for k-fold CV is  $k$ , which defines how many folds the dataset will be split into. Hence, as shown in Table 7, different folds (3, 5, and 10) were used in this study. When these k-fold testing values are compared, it is possible to conclude that the CNN model provides the most accurate results with  $k=5$  for all stations. With the three ML models, estimated values relating to monthly Ep have been plotted against the measured values for each station as shown in Figs. 12, 13, 14 and 15. The lower-level pertaining to scatter plot and an improved fit with respect to the estimated data with that of the values observed in the 1:1 line are the clear indicators suggesting the superiority with respect to the CNN model compared to other models. Even though Figs. 12, 13, 14 and 15 as well as Table 6 display the observed and estimated values for all the models, and also the evaluation criteria, the Taylor diagram (TD) was employed to compare the methods presented in this research. The primary concept of the TD is to represent the closest prediction model with actual corresponding observation in the 2-D scaling

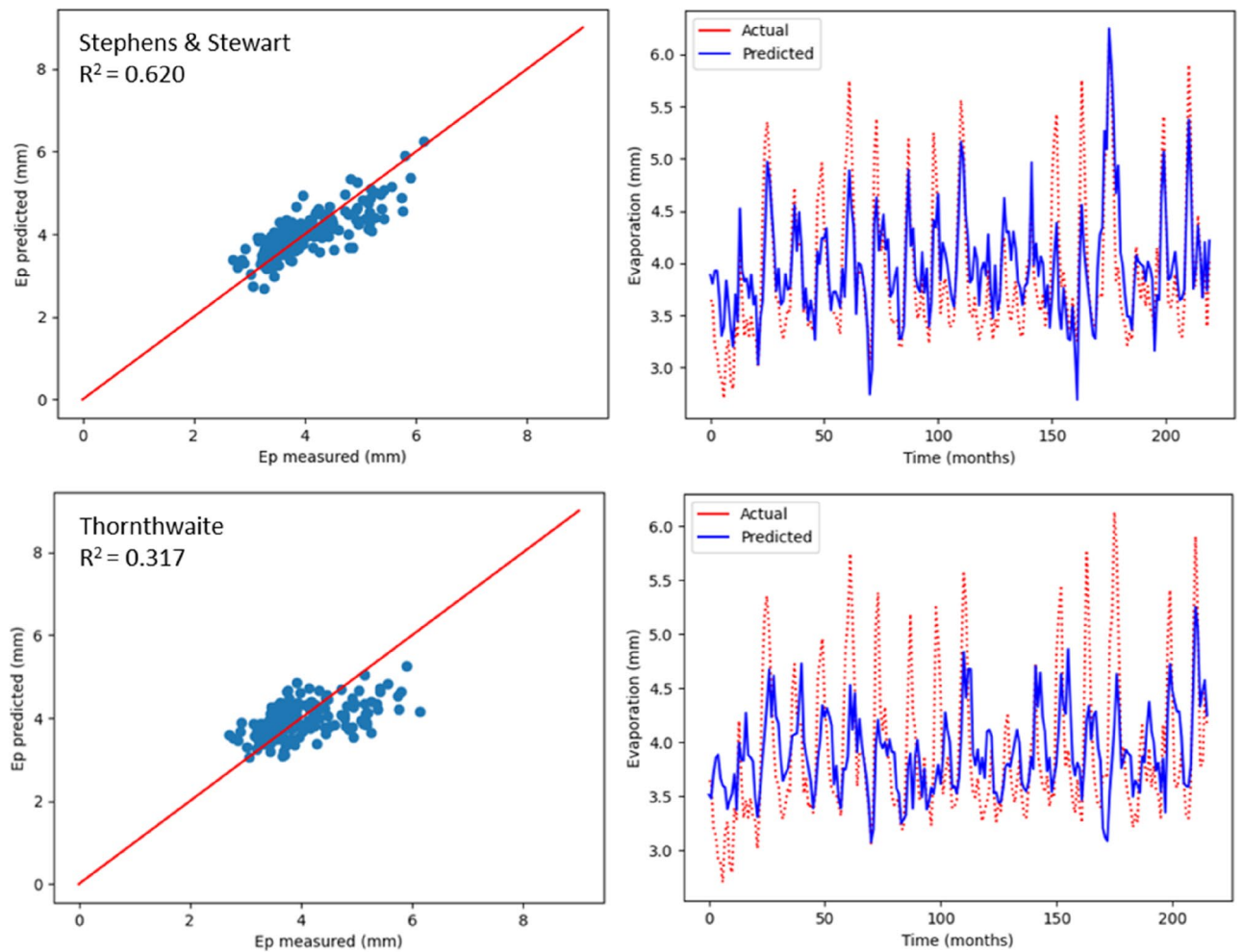


**Figure 7.** The process of developing a prediction model.

Station	Model	R <sup>2</sup>	MAE	MSE	RMSE	RAE	NSE
Bayan Lepas	Stephens & Stewart	0.620	0.328	0.167	0.409	0.631	0.621
	Thornthwaite	0.317	0.431	0.306	0.553	0.820	0.317
Ipoh	Stephens & Stewart	0.649	0.231	0.085	0.292	0.585	0.650
	Thornthwaite	0.635	0.243	0.088	0.296	0.615	0.636
KLIA Sepang	Stephens & Stewart	0.580	0.244	0.098	0.314	0.670	0.581
	Thornthwaite	0.256	0.325	0.175	0.418	0.891	0.257
Kuantan	Stephens & Stewart	0.696	0.245	0.085	0.292	0.572	0.697
	Thornthwaite	0.497	0.284	0.144	0.380	0.657	0.498

**Table 5.** Statistical results of Stephens & Stewart and Thornthwaite empirical models for prediction Ep at Bayan Lepas, Ipoh, KLIA Sepang and Kuantan stations.

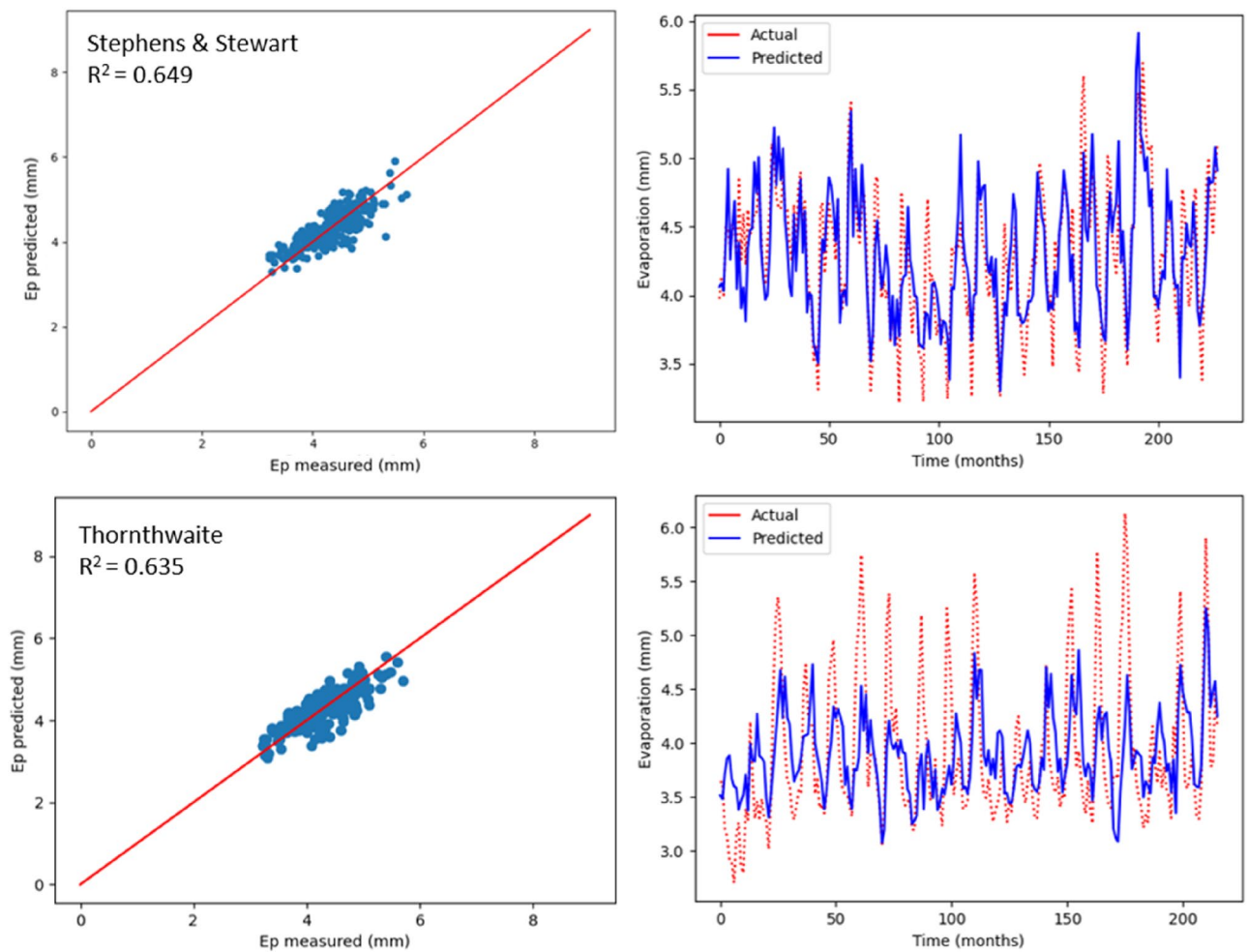
(correlation coefficient on polar axis and standard deviation on radial axis). Standard deviation is with respect to how much, on average, measurements vary from each other. Thus, the relative value of SDP from SDA indicates the level of accuracy. The value of SDP from S.D.A. pertains to lower accuracy. Greater difference refers to lower precision. Therefore, in Fig. 16, it can be noticed that the CNN-9 was better compared to other methodologies, which had SD of 0.65 closer to the actual SD of 0.66 in Bayan Lepas, SD of 0.47 to the actual SD of 0.49 in Ipoh, SD of 0.47 to the actual SD of 0.48 in KLIA Sepang, and SD of 0.52 to the actual SD of 0.53 in Kuantan. The comparison of predicted and actual Ep monthly values generated by the most exact models is displayed in Fig. 16, which demonstrated that the ML models are superior to other models generally, while the CNN-9 is superior to the ML models in particular.



**Figure 8.** Scatter plot of measured Ep versus predicted Ep for the proposed empirical models for Bayan Lepas station.

As per Table 6, realisation of the best prediction accuracy was possible through the models employing the complete meteorological dataset ( $T_{\max}$ ,  $T_{\min}$ ,  $R_s$ ,  $S_w$ , RH and Ep) with regards to all stations, when compared with combinations pertaining to other incomplete data input. This showed that the model prediction's accuracy improved in general with additional input parameters, which was similar to the results seen in the earlier studies<sup>3,34</sup>. Four input parameters that have not included  $R_s$  or  $S_w$  were adequate to achieve acceptable accuracy with regards to estimation of monthly Ep. When only mean temperature data were available, ML models, including the CNN model, were found to be insufficient for all stations. This implied that employing the powerful capabilities, such as AI may not improve the ML model prediction accuracy, particularly when meteorological inputs are restricted. Besides, with regards to all ML models, the prediction accuracy improved slightly by using Ep as an input. However, the statistical values with regards to machine learning models were close to complete meteorological inputs (i.e., using Ep as an input) by employing the input combination pertaining to  $T_{\min}$ ,  $T_{\max}$ ,  $S_w$ ,  $R_s$  and RH. This suggested that the estimated monthly Ep values through machine learning models were in general in line with those of the measured monthly Ep values.

Apart from the robustness and convenience associated with DLs automated feature extraction, it was seen that the proposed deep learning models consistently outdid the RF model when it comes to prediction of Ep. Thus, these research results were in line with the previous studies<sup>53,75</sup>, which mentioned deep learning to be a powerful modelling technique that allows learning the complex and non-linear behaviours pertaining to evaporation. Particularly, it was seen that the CNN model was better than other DL models, such as DNN, which indicates the CNN model's high potential when it comes to modelling and mapping evaporation when it is difficult for most of the ML models. The effectiveness pertaining to CNN in capturing and analysing the non-linearity and complexity behaviours of evaporation with greater efficacy could be due to the convolutional characteristic of 1D-CNN, i.e., a large number of convolutional kernels are applied by CNN to the inputs for extracting information extensively, which is helpful for time series forecast. However, DLs versus RFs need to be compared carefully, since there is a chance of underestimating the capacity of RFs when special consideration is not given. Thus, the time needed to run and tune the models also needs to be considered when objectively comparing between DL and RF models. Although training time can be influenced by several factors (e.g., model complexity, number of



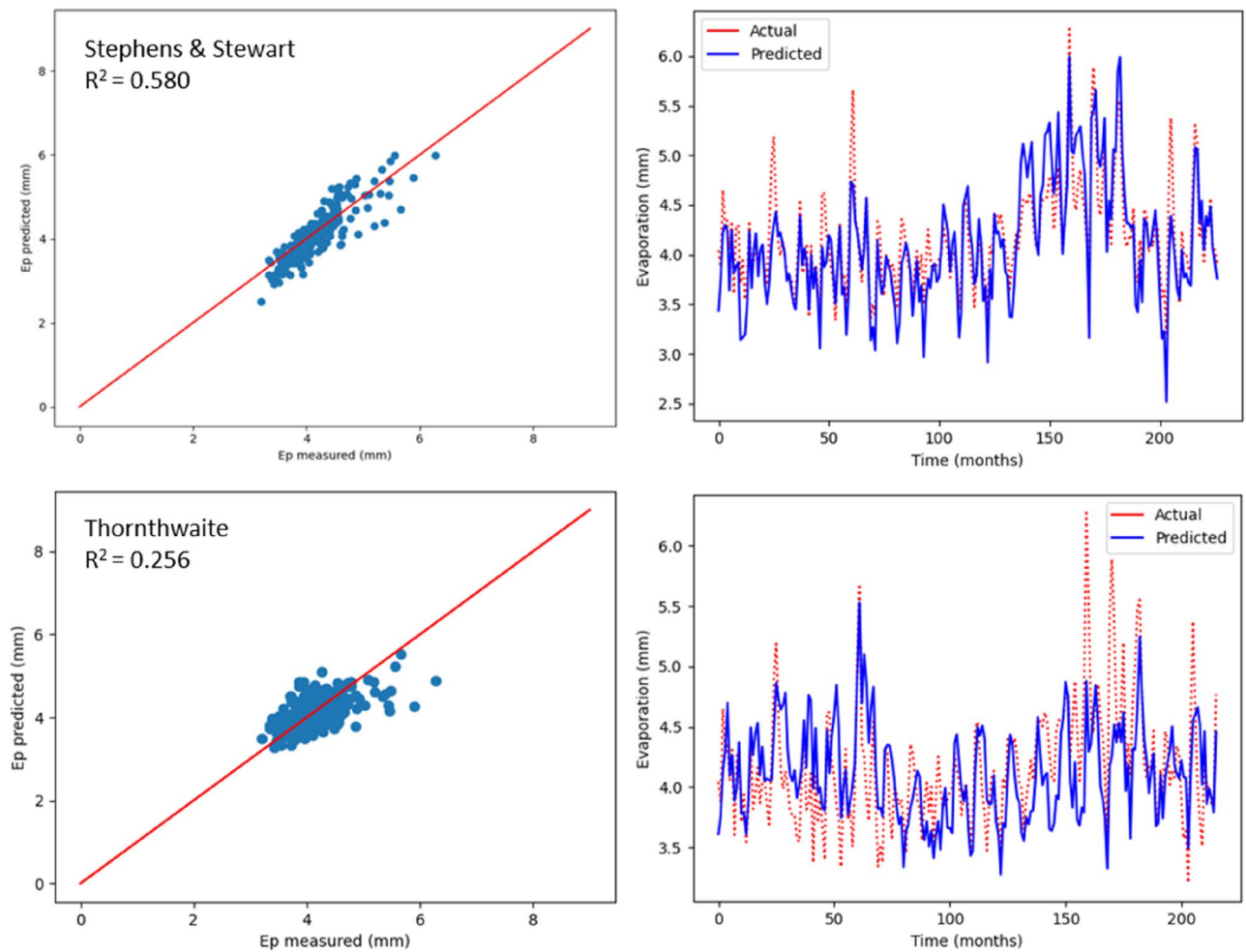
**Figure 9.** Scatter plot of measured Ep versus predicted Ep for the proposed empirical models for Ipoh station.

inputs employed), in general, RF has been found to be faster in tuning and training versus DL. The application of DL includes training time as one of the challenges. In addition to this, it is challenging to optimise DL since no formula has been identified that can guarantee converging of DL to a good solution. Moreover, when compared with the RF, larger data sets are required for DL to learn the evaporation properties. Due to this, even though deep learning is regarded to be very powerful when it comes to capturing complex and non-linear behaviours, there exist certain challenges that need to be taken into account when constructing deep learning prediction models.

With regards to the above statement, the CNN model was seen to be able to model pan evaporation with high prediction accuracy. However, for validating the developed predictive model's predictability, a comparison was performed for the results pertaining to the current study versus other AI models exposed to same climatic conditions. Mustafa et al.<sup>76</sup> reported an  $R^2$  value of 0.97 with regards to their best-performing SVM model during validation period by employing the Support Vector Machine (SVM) method in the Ipoh region based on the same data as used in the current research, versus an  $R^2$  value of 0.98 that was identified in the current study. It was also seen that the CNN model was better compared to other AI methods, including K-Nearest Neighbours (KNN), which was recently used in the Ipoh region based on the same data as employed in the present study (M.A, M.A.I, A.N.A, and Y.F.H). Satisfactory performance was reported by applying the KNN, which gave an  $R^2$  value of 0.94. Based on this, the study concluded that DL in general, and CNN in particular, can be used as optimistic predictive models in hydrological applications such as evaporation due to the excellent features described earlier. Moreover, investigation will be carried out with regards to the application of the proposed methodology for different regions throughout Malaysia by employing different data sets in order to construct a reliable generalised model for evaporation prediction.

**Comparison of empirical and ML models.** Table 8 demonstrates the performances for two empirical models to perform prediction of monthly Ep, which are then compared to their respective ML models using same input combinations for Bayan Lepas, Ipoh, KLIA Sepang and Kuantan weather stations. As an initial observation, with regards to input combination of  $R_s$  and  $T_a$  for all stations, the radiation-based model (Stewart and Stephens) offered the lowest prediction accuracy ( $R^2$  values: 0.620, 0.649, 0.580, and 0.696) in comparison with all ML models. On the other hand, the machine learning models (i.e., RF-1, CNN-1 and DNN-1) were seen to



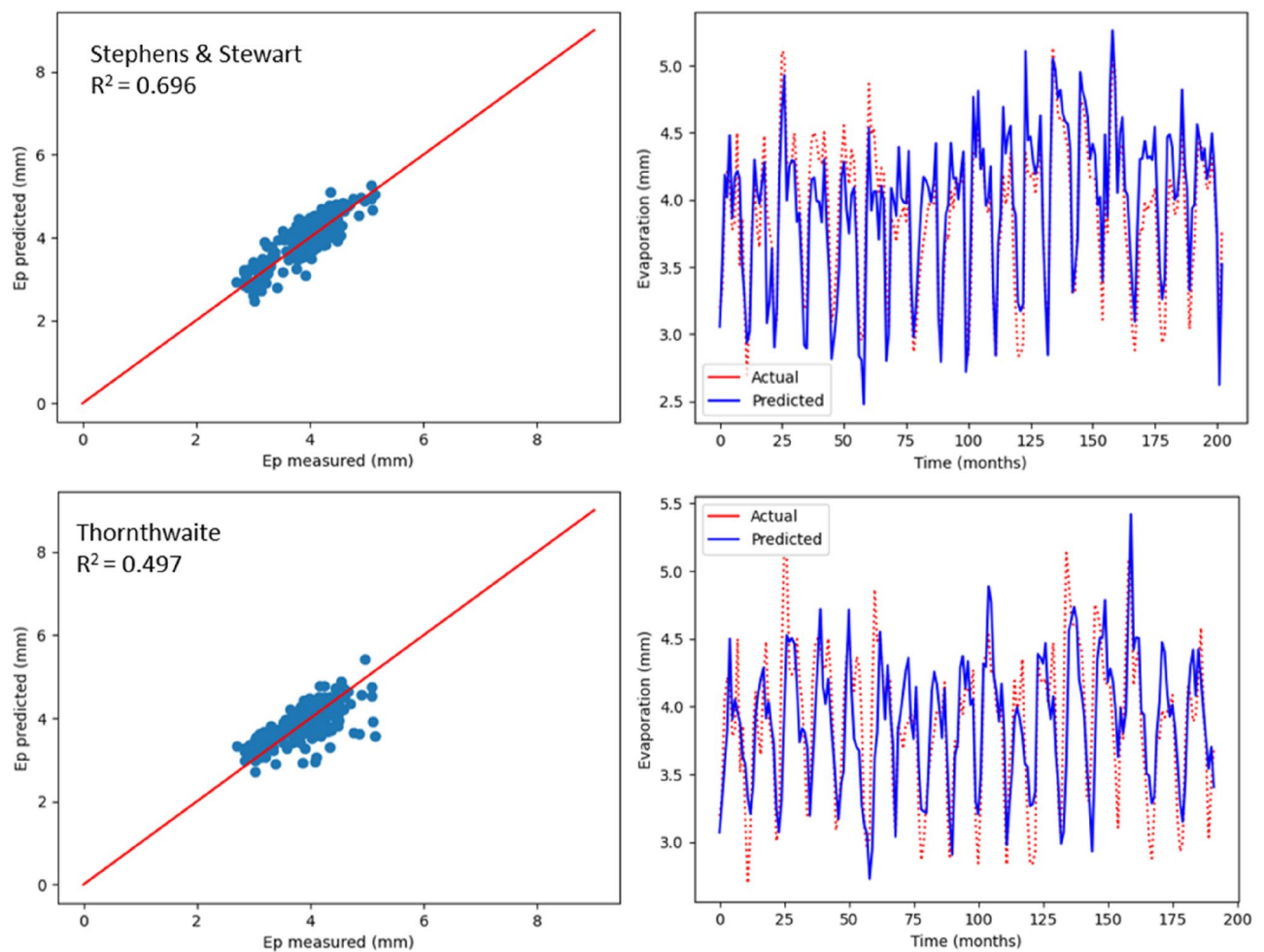


**Figure 10.** Scatter plot of measured  $E_p$  versus predicted  $E_p$  for the proposed empirical models for KLIA Sepang station.

perform excellently to achieve high prediction accuracy versus the temperature-based model (Thornthwaite) based on the input combination of just  $T_a$ . Based on the statistical results presented in Table 7, the higher performance of ML models was evident versus empirical models, and could also considerably enhance the prediction accuracy of monthly  $E_p$  even when employing the same input parameters, depending on their superior capabilities to carry out non-linear and complex tasks. Furthermore, it has been seen that higher accuracy was achieved with the deep learning models (i.e., DNN and CNN) in terms of forecasting evaporation versus the tree-based model (i.e., RF). This can be attributed to the deep learning feature catching concealed properties, which signifies that deep learning can be regarded as more powerful approach for predicting evaporation. In this regard, although the RF was seen to marginally outperform the DL models for few cases, it is evident that this is a single case since the DL models are regarded to be more consistent and could also offer higher accuracy versus empirical and tree-based methods based on all the different input sets at all stations.

## Conclusion

This study is conducted to determine the monthly  $E_p$  losses by employing RF, DNN, and CNN techniques. Monthly data from four weather stations in Malaysia were employed to assess the capabilities of the three AI approaches in predicting the  $E_p$  rates. Time series data pertaining to monthly  $E_p$ , such as  $T_{max}$ ,  $T_{min}$ ,  $T_a$ , RH,  $S_w$ ,  $R_s$ , and  $E_p$ , between the years 2000–2019 were used to set up the evaluated models. The data was divided into two parts: 20% for testing (validation) and 80% for training (calibration). The PCC values were used to select the input parameters (predictors) in order to identify the most effective input combinations for ML models. The developed ML models were compared to two empirical models, one is temperature-based model (Thornthwaite)



**Figure 11.** Scatter plot of measured Ep versus predicted Ep for the proposed empirical models for Kuantan station.

while the other is radiation-based model (Stephens & Stewart). Standard statistical measures were employed to assess the performance of each model as well as their effectiveness pertaining to evaporation forecasting. Furthermore, the accuracy of the studied models was evaluated using the Taylor diagram. The investigation yielded the following results:

- The three developed ML models were found to outperform the empirical methods and to significantly improve the precision of monthly Ep estimates even when using the same combinations of inputs.
- Both RF and DL methods can accurately predict the monthly Ep. In particular, when it comes to predicting Ep, the DL approach (i.e., CNN and DNN) was found to slightly outperform the RF model.
- The best ML prediction accuracy could be achieved with models that employed complete meteorological datasets ( $T_{max}$ ,  $T_{min}$ ,  $R_s$ ,  $S_w$ , RH and Ep) with regards to all stations, when compared with other combinations of incomplete data input.
- As seen in the results, the monthly evaporation losses can be successfully modelled based on the CNN structure along with enhanced accuracy versus other models that were accounted in this study. Moreover, estimation results based on the CNN model were seen to outdo versus other AI approaches that were studied in the same regions by employing the same data.
- In the future, the applicability of the proposed methodology to different regions in Malaysia can be assessed using different data sets with the aim of building a dependable generalised model for predicting evaporation.

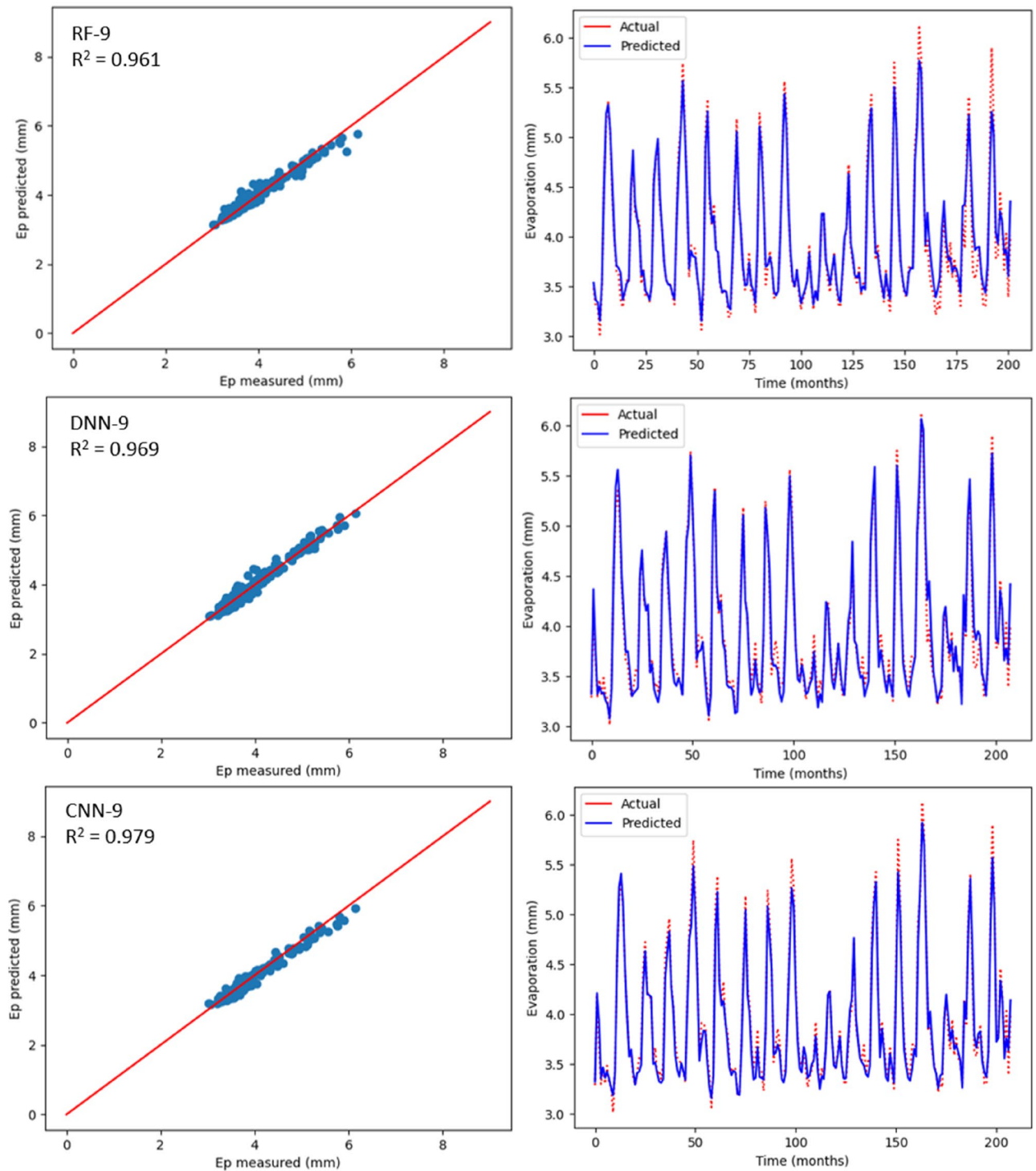
Station/model	R <sup>2</sup>	MAE	MSE	RMSE	RAE	NSE
<b>Bayan Lepas</b>						
RF-1	0.562	0.324	0.195	0.441	0.619	0.563
RF-2	0.727	0.248	0.121	0.348	0.474	0.727
RF-3	0.797	0.219	0.090	0.300	0.418	0.798
RF-4	0.848	0.197	0.066	0.258	0.377	0.849
RF-5	0.857	0.186	0.062	0.250	0.357	0.858
RF-6	0.909	0.149	0.039	0.198	0.288	0.910
RF-7	0.868	0.174	0.057	0.238	0.335	0.869
RF-8	0.953	0.115	0.020	0.142	0.222	0.953
RF-9	<b>0.961</b>	<b>0.088</b>	<b>0.016</b>	<b>0.128</b>	<b>0.169</b>	<b>0.962</b>
DNN-1	0.492	0.351	0.226	0.475	0.671	0.493
DNN-2	0.800	0.224	0.088	0.298	0.429	0.801
DNN-3	0.807	0.223	0.085	0.292	0.427	0.808
DNN-4	0.857	0.188	0.063	0.251	0.359	0.858
DNN-5	0.838	0.197	0.070	0.266	0.378	0.839
DNN-6	0.908	0.147	0.039	0.199	0.283	0.909
DNN-7	0.878	0.158	0.052	0.229	0.304	0.879
DNN-8	0.929	0.135	0.030	0.174	0.260	0.930
DNN-9	<b>0.969</b>	<b>0.087</b>	<b>0.012</b>	<b>0.113</b>	<b>0.169</b>	<b>0.970</b>
CNN-1	0.623	0.290	0.167	0.409	0.554	0.624
CNN-2	0.767	0.231	0.103	0.321	0.442	0.768
CNN-3	0.810	0.222	0.084	0.290	0.425	0.811
CNN-4	0.868	0.178	0.058	0.240	0.341	0.869
CNN-5	0.882	0.167	0.051	0.227	0.321	0.883
CNN-6	0.911	0.148	0.038	0.195	0.286	0.912
CNN-7	0.890	0.174	0.047	0.217	0.335	0.891
CNN-8	0.965	0.084	0.015	0.122	0.163	0.966
CNN-9	<b>0.979</b>	<b>0.071</b>	<b>0.008</b>	<b>0.092</b>	<b>0.138</b>	<b>0.980</b>
<b>Ipoh</b>						
RF-1	0.652	0.227	0.084	0.290	0.573	0.653
RF-2	0.737	0.194	0.064	0.253	0.489	0.738
RF-3	0.478	0.291	0.126	0.356	0.735	0.479
RF-4	0.720	0.203	0.068	0.261	0.511	0.721
RF-5	0.734	0.197	0.065	0.256	0.494	0.735
RF-6	0.865	0.140	0.033	0.182	0.348	0.866
RF-7	0.832	0.146	0.041	0.204	0.365	0.833
RF-8	0.901	0.111	0.024	0.157	0.276	0.902
RF-9	<b>0.959</b>	<b>0.066</b>	<b>0.010</b>	<b>0.101</b>	<b>0.163</b>	<b>0.960</b>
DNN-1	0.719	0.205	0.068	0.261	0.519	0.719
DNN-2	0.793	0.162	0.050	0.224	0.408	0.794
DNN-3	0.511	0.277	0.119	0.344	0.701	0.521
DNN-4	0.800	0.173	0.048	0.221	0.436	0.801
DNN-5	0.804	0.170	0.048	0.219	0.424	0.805
DNN-6	0.903	0.123	0.024	0.155	0.307	0.904
DNN-7	0.882	0.122	0.029	0.171	0.305	0.883
DNN-8	0.922	0.106	0.019	0.139	0.264	0.923
DNN-9	<b>0.945</b>	<b>0.094</b>	<b>0.013</b>	<b>0.116</b>	<b>0.232</b>	<b>0.946</b>
CNN-1	0.749	0.192	0.061	0.247	0.486	0.750
CNN-2	0.831	0.152	0.041	0.203	0.382	0.831
CNN-3	0.550	0.270	0.109	0.331	0.683	0.550
CNN-4	0.811	0.171	0.046	0.214	0.430	0.812
CNN-5	0.837	0.157	0.040	0.200	0.393	0.838
CNN-6	0.935	0.096	0.016	0.126	0.239	0.936
CNN-7	0.904	0.113	0.023	0.154	0.281	0.905
CNN-8	0.959	0.071	0.010	0.100	0.176	0.960
CNN-9	<b>0.980</b>	<b>0.053</b>	<b>0.004</b>	<b>0.069</b>	<b>0.132</b>	<b>0.981</b>
<b>KLIA Sepang</b>						
RF-1	0.293	0.299	0.167	0.409	0.817	0.294
Continued						

Station/model	R <sup>2</sup>	MAE	MSE	RMSE	RAE	NSE
RF-2	0.730	0.171	0.064	0.253	0.467	0.731
RF-3	0.819	0.077	0.042	0.207	0.212	0.819
RF-4	0.878	0.066	0.028	0.169	0.180	0.879
RF-5	0.880	0.120	0.028	0.168	0.328	0.881
RF-6	0.910	0.108	0.021	0.146	0.293	0.911
RF-7	0.909	0.109	0.021	0.147	0.296	0.910
RF-8	0.931	0.093	0.016	0.128	0.255	0.932
RF-9	<b>0.943</b>	<b>0.070</b>	<b>0.013</b>	<b>0.117</b>	<b>0.190</b>	<b>0.944</b>
DNN-1	0.358	0.287	0.152	0.390	0.783	0.359
DNN-2	0.700	0.183	0.071	0.267	0.499	0.701
DNN-3	0.820	0.141	0.042	0.206	0.385	0.821
DNN-4	0.859	0.129	0.033	0.182	0.355	0.860
DNN-5	0.893	0.059	0.025	0.159	0.161	0.894
DNN-6	0.924	0.101	0.018	0.134	0.275	0.925
DNN-7	0.891	0.119	0.026	0.161	0.324	0.892
DNN-8	0.939	0.100	0.014	0.121	0.269	0.940
DNN-9	<b>0.959</b>	<b>0.071</b>	<b>0.009</b>	<b>0.099</b>	<b>0.191</b>	<b>0.960</b>
CNN-1	0.375	0.284	0.148	0.384	0.774	0.376
CNN-2	0.717	0.178	0.067	0.259	0.485	0.718
CNN-3	0.827	0.079	0.040	0.202	0.216	0.827
CNN-4	0.830	0.137	0.040	0.200	0.375	0.831
CNN-5	0.899	0.114	0.024	0.155	0.313	0.900
CNN-6	0.921	0.101	0.019	0.137	0.274	0.922
CNN-7	0.919	0.101	0.019	0.139	0.274	0.920
CNN-8	0.949	0.081	0.012	0.110	0.219	0.950
CNN-9	<b>0.965</b>	<b>0.079</b>	<b>0.008</b>	<b>0.091</b>	<b>0.214</b>	<b>0.966</b>
<b>Kuantan</b>						
RF-1	0.647	0.241	0.098	0.314	0.567	0.648
RF-2	0.811	0.197	0.054	0.232	0.458	0.811
RF-3	0.688	0.237	0.087	0.295	0.556	0.689
RF-4	0.852	0.157	0.041	0.203	0.370	0.853
RF-5	0.888	0.059	0.005	0.074	0.330	0.889
RF-6	0.904	0.122	0.027	0.165	0.283	0.905
RF-7	0.889	0.138	0.031	0.178	0.321	0.890
RF-8	0.930	0.108	0.019	0.140	0.253	0.931
RF-9	<b>0.956</b>	<b>0.088</b>	<b>0.012</b>	<b>0.110</b>	<b>0.207</b>	<b>0.957</b>
DNN-1	0.649	0.248	0.098	0.314	0.581	0.650
DNN-2	0.814	0.180	0.052	0.229	0.422	0.815
DNN-3	0.709	0.228	0.081	0.285	0.573	0.710
DNN-4	0.843	0.166	0.044	0.210	0.389	0.844
DNN-5	0.876	0.146	0.035	0.187	0.339	0.877
DNN-6	0.917	0.121	0.023	0.154	0.282	0.917
DNN-7	0.898	0.132	0.028	0.168	0.309	0.898
DNN-8	0.944	0.103	0.015	0.126	0.238	0.945
DNN-9	<b>0.958</b>	<b>0.084</b>	<b>0.011</b>	<b>0.108</b>	<b>0.198</b>	<b>0.959</b>
CNN-1	0.654	0.239	0.096	0.310	0.562	0.655
CNN-2	0.816	0.187	0.051	0.227	0.437	0.817
CNN-3	0.747	0.210	0.071	0.266	0.491	0.748
CNN-4	0.871	0.149	0.036	0.190	0.350	0.872
CNN-5	0.877	0.143	0.034	0.185	0.335	0.878
CNN-6	0.906	0.124	0.026	0.163	0.287	0.907
CNN-7	0.897	0.130	0.028	0.169	0.304	0.898
CNN-8	0.956	0.086	0.012	0.112	0.198	0.956
CNN-9	<b>0.962</b>	<b>0.084</b>	<b>0.010</b>	<b>0.103</b>	<b>0.198</b>	<b>0.962</b>

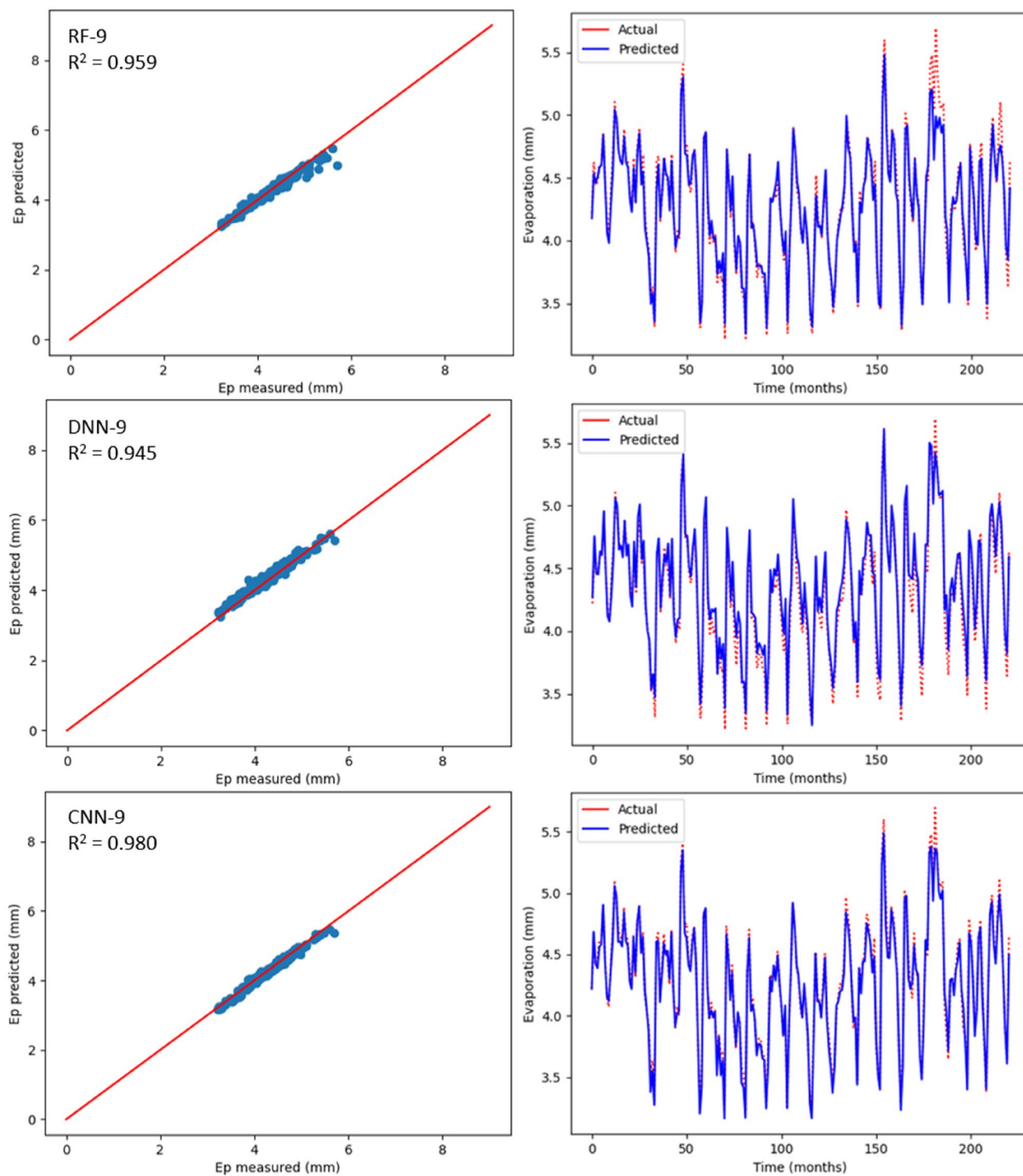
**Table 6.** Statistical results (testing period) of the three machine learning models for predicting monthly Ep under nine input combinations of meteorological variables for Bayan Lepas, Ipoh, KLIA Sepang and Kuantan stations. Bold indicates the optimum statistical parameters.

k-fold CV	R-Squared values			
	Bayan Lepas	Ipoh	KLIA Sepang	Kuantan
K=3	0.94	0.96	0.95	0.94
K=5	0.97	0.98	0.96	0.96
K=10	0.96	0.97	0.95	0.95

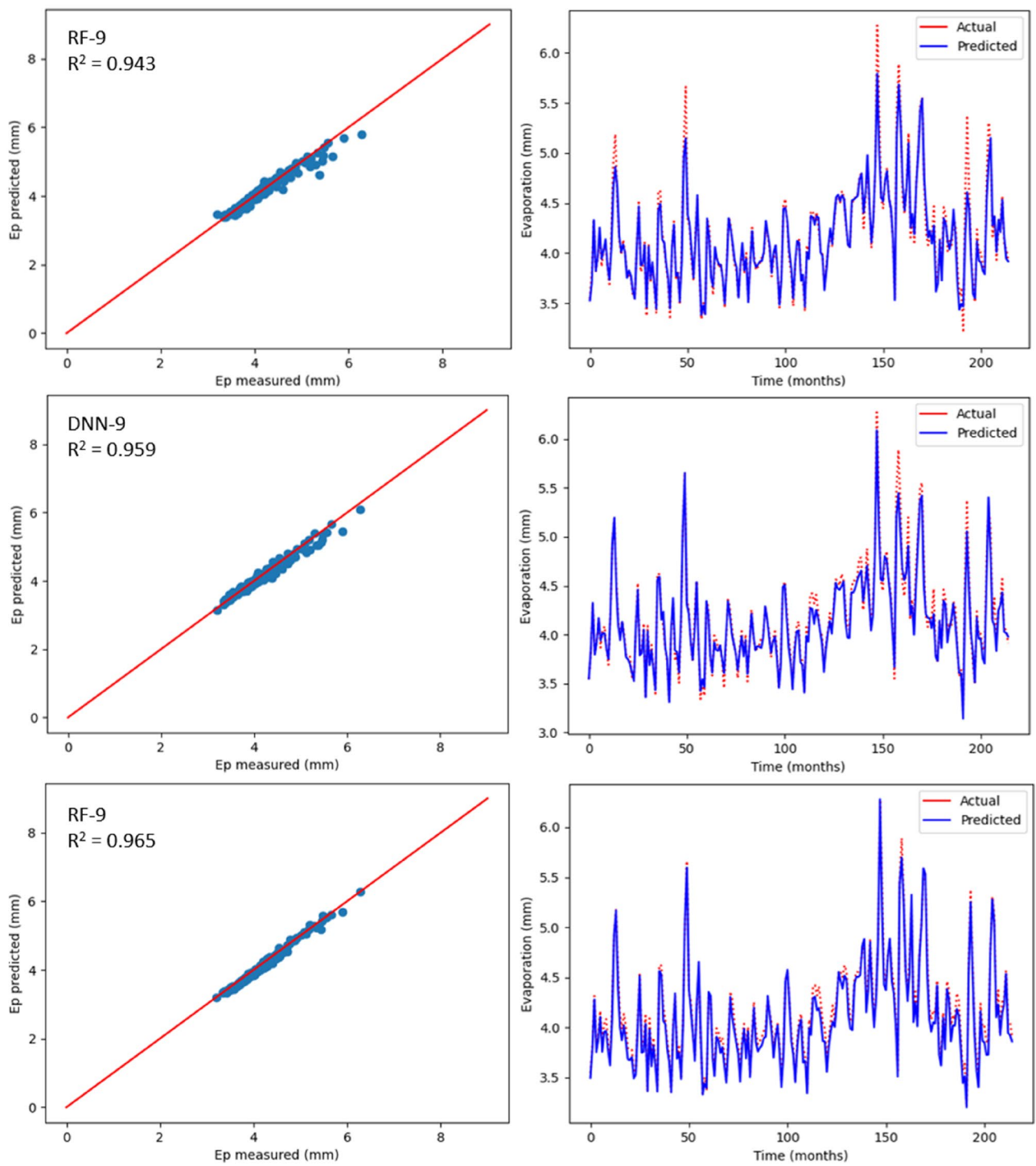
**Table 7.** Time series cross-validation.



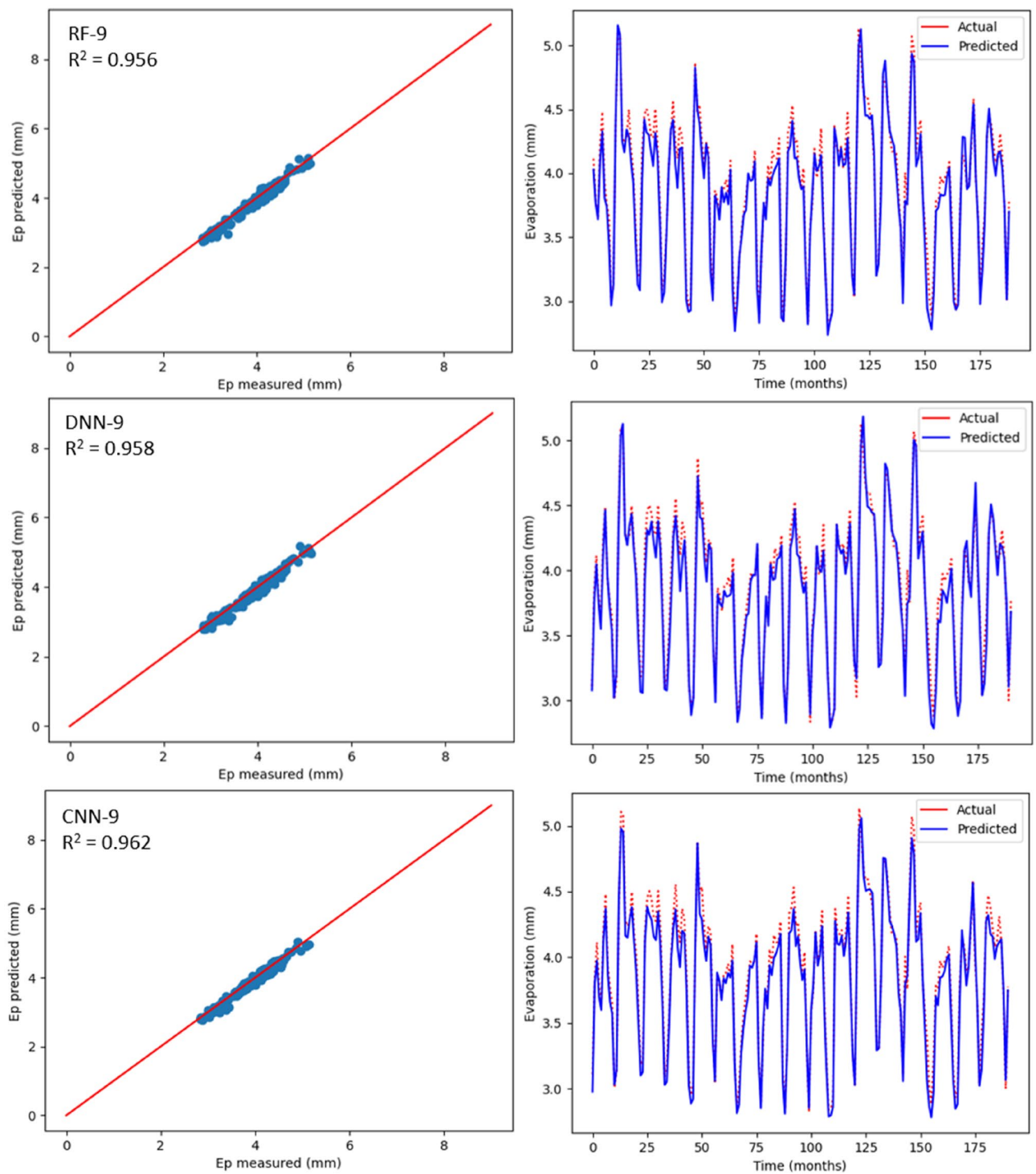
**Figure 12.** Scatter plot of measured Ep versus predicted Ep for the proposed machine learning models for Bayan Lepas station.



**Figure 13.** Scatter plot of measured Ep versus predicted Ep for the proposed machine learning models for Ipoh station.



**Figure 14.** Scatter plot of measured Ep versus predicted Ep for the proposed machine learning models for KLIA Sepang station.



**Figure 15.** Scatter plot of measured Ep versus predicted Ep for the proposed machine learning models for Kuantan station.



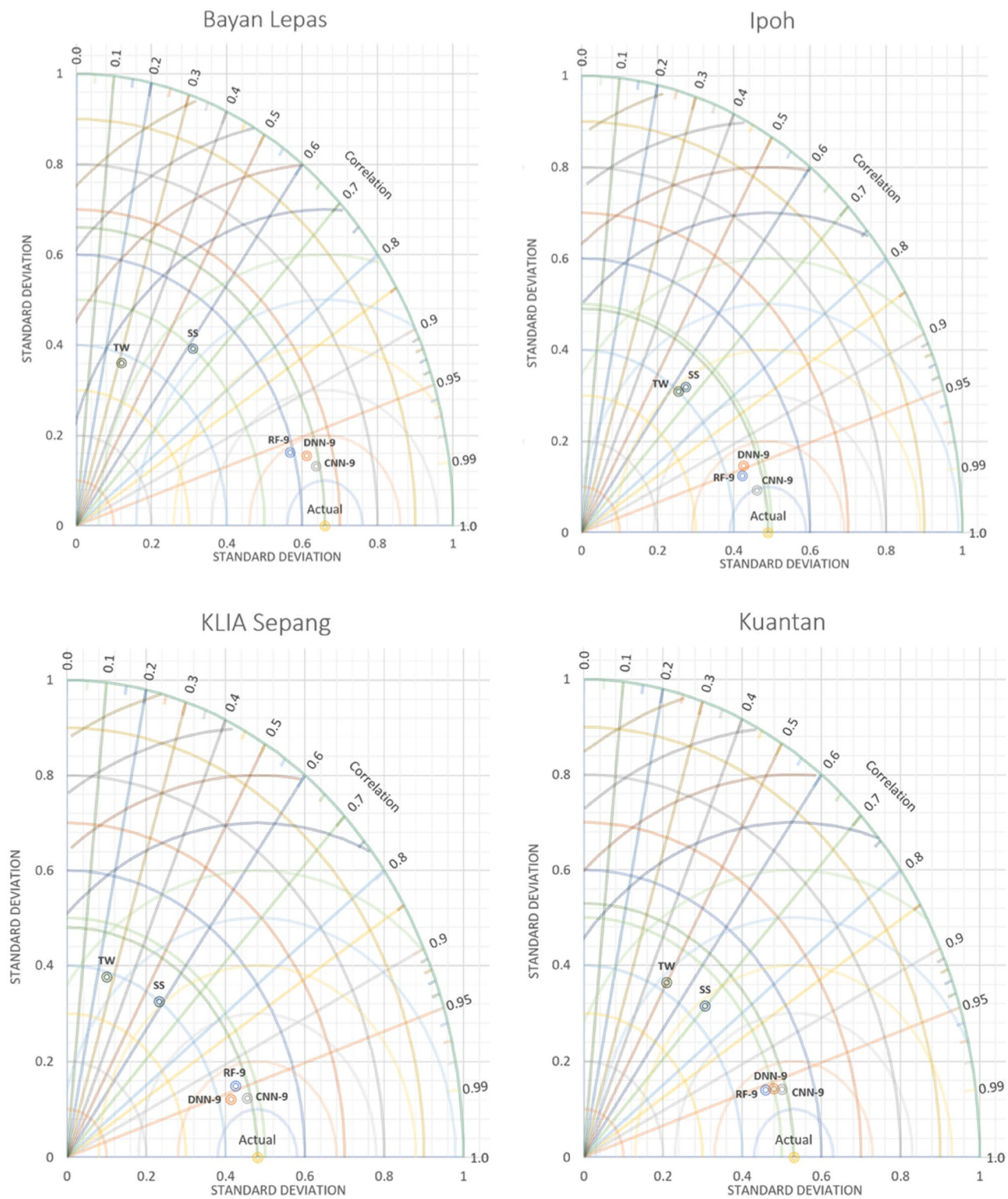


Figure 16. Taylor diagram of predicted monthly pan evaporation during the validation stage.

Input combination	Station/model	R <sup>2</sup>	MAE	MSE	RMSE	RAE	NSE
<b>Bayan Lepas</b>							
T <sub>a</sub> , R <sub>s</sub>	Stephens & Stewart	0.620	0.328	0.167	0.409	0.631	0.621
	RF-2	0.727	0.248	0.121	0.348	0.474	0.727
	DNN-2	0.800	0.224	0.088	0.298	0.429	0.801
	CNN-2	0.767	0.231	0.103	0.321	0.442	0.768
T <sub>a</sub>	Thornthwaite	0.317	0.431	0.306	0.553	0.820	0.317
	RF-1	0.562	0.324	0.195	0.441	0.619	0.563
	DNN-1	0.492	0.351	0.226	0.475	0.671	0.493
	CNN-1	0.623	0.290	0.167	0.409	0.554	0.624
	<b>Ipoh</b>						
T <sub>a</sub> , R <sub>s</sub>	Stephens & Stewart	0.649	0.231	0.085	0.292	0.585	0.650
	RF-2	0.737	0.194	0.064	0.253	0.489	0.738
	DNN-2	0.793	0.162	0.050	0.224	0.408	0.794
	CNN-2	0.831	0.152	0.041	0.203	0.382	0.831
T <sub>a</sub>	Thornthwaite	0.635	0.243	0.088	0.296	0.615	0.636
	RF-1	0.652	0.227	0.084	0.290	0.573	0.653
	DNN-1	0.719	0.205	0.068	0.261	0.519	0.719
	CNN-1	0.749	0.192	0.061	0.247	0.486	0.750
	<b>KLIA Sepang</b>						
T <sub>a</sub> , R <sub>s</sub>	Stephens & Stewart	0.580	0.244	0.098	0.314	0.670	0.581
	RF-2	0.730	0.171	0.064	0.253	0.467	0.731
	DNN-2	0.700	0.183	0.071	0.267	0.499	0.701
	CNN-2	0.717	0.178	0.067	0.259	0.485	0.718
T <sub>a</sub>	Thornthwaite	0.256	0.325	0.175	0.418	0.891	0.257
	RF-1	0.293	0.299	0.167	0.409	0.817	0.294
	DNN-1	0.358	0.287	0.152	0.390	0.783	0.359
	CNN-1	0.375	0.284	0.148	0.384	0.774	0.376
	<b>Kuantan</b>						
T <sub>a</sub> , R <sub>s</sub>	Stephens & Stewart	0.696	0.245	0.085	0.292	0.572	0.697
	RF-2	0.811	0.197	0.054	0.232	0.458	0.811
	DNN-2	0.814	0.180	0.052	0.229	0.422	0.815
	CNN-2	0.816	0.187	0.051	0.227	0.437	0.817
T <sub>a</sub>	Thornthwaite	0.497	0.284	0.144	0.380	0.657	0.498
	RF-1	0.647	0.241	0.098	0.314	0.567	0.648
	DNN-1	0.649	0.248	0.098	0.314	0.581	0.650
	CNN-1	0.654	0.239	0.096	0.310	0.562	0.655

**Table 8.** Statistical results of the empirical and machine learning models under the same input combination for Bayan Lepas, Ipoh, KLIA Sepang and Kuantan weather stations.

## Data availability

The datasets used during the current study are available from the first author on reasonable request.

Received: 21 February 2022; Accepted: 22 July 2022

Published online: 30 July 2022

## References

- Shiri, J. & Kişi, Ö. Application of artificial intelligence to estimate daily pan evaporation using available and estimated climatic data in the Khozestan Province (South Western Iran). *J. Irrig. Drain. Eng.* **137**, 412–425 (2011).
- Gundalia, M. J. & Dholakia, M. B. Estimation of pan evaporation using mean air temperature and radiation for monsoon season in Junagadh region. *Int. J. Eng. Res. Appl.* **3**, 64–70 (2013).
- Fan, J., Wu, L., Zhang, F., Xiang, Y. & Zheng, J. Climate change effects on reference crop evapotranspiration across different climatic zones of China during 1956–2015. *J. Hydrol.* **542**, 923–937 (2016).
- Feng, Y., Jia, Y., Zhang, Q., Gong, D. & Cui, N. National-scale assessment of pan evaporation models across different climatic zones of China. *J. Hydrol.* **564**, 314–328 (2018).
- Chen, Y., He, L., Li, J. & Zhang, S. Multi-criteria design of shale-gas-water supply chains and production systems towards optimal life cycle economics and greenhouse gas emissions under uncertainty. *Comput. Chem. Eng.* **109**, 216–235 (2018).
- Wang, Q., Wang, W., Zhong, Z., Wang, H. & Fu, Y. Variation in glomalin in soil profiles and its association with climatic conditions, shelterbelt characteristics, and soil properties in poplar shelterbelts of Northeast China. *J. For. Res.* **31**, 279–290 (2020).
- Allawi, M. F. *et al.* Reservoir evaporation prediction modeling based on artificial intelligence methods. *Water* **11**, 1226 (2019).

8. Morton, F. I. *Evaporation and Climate: A Study in Cause and Effect, Scientific Series No. 4* (International Water Branch, Department Energy, Mines Resource, 1968).
9. Vicente-Serrano, S. M. *et al.* A comparison of temporal variability of observed and model-based pan evaporation over Uruguay (1973–2014). *Int. J. Climatol.* **38**, 337–350 (2018).
10. Fan, J. *et al.* Evaluation and development of temperature-based empirical models for estimating daily global solar radiation in humid regions. *Energy* **144**, 903–914 (2018).
11. Wu, L. *et al.* Hybrid extreme learning machine with meta-heuristic algorithms for monthly pan evaporation prediction. *Comput. Electron. Agric.* **168**, 105115 (2020).
12. Keshtegar, B., Piri, J. & Kisi, O. A nonlinear mathematical modeling of daily pan evaporation based on conjugate gradient method. *Comput. Electron. Agric.* **127**, 120–130 (2016).
13. Kişi, Ö. Daily pan evaporation modelling using multi-layer perceptrons and radial basis neural networks. *Hydrol. Process. Int. J.* **23**, 213–223 (2009).
14. Arunkumar, R. & Jothiprakash, V. Reservoir evaporation prediction using data-driven techniques. *J. Hydrol. Eng.* **18**, 40–49 (2013).
15. Quinn, R., Parker, A. & Rushton, K. Evaporation from bare soil: Lysimeter experiments in sand dams interpreted using conceptual and numerical models. *J. Hydrol.* **564**, 909–915 (2018).
16. Moeeni, H. & Bonakdari, H. Impact of normalization and input on ARMAX-ANN model performance in suspended sediment load prediction. *Water Resour. Manage.* **32**, 845–863 (2018).
17. Myronidis, D., Ioannou, K., Fotakis, D. & Dörfinger, G. Streamflow and hydrological drought trend analysis and forecasting in Cyprus. *Water Resour. Manage.* **32**, 1759–1776 (2018).
18. Wang, W., Chau, K., Xu, D. & Chen, X.-Y. Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition. *Water Resour. Manage.* **29**, 2655–2675 (2015).
19. Ghamariadayan, M. & Imteaz, M. A. A wavelet artificial neural network method for medium-term rainfall prediction in Queensland (Australia) and the comparisons with conventional methods. *Int. J. Climatol.* **41**, E1396–E1416 (2021).
20. Aziz, M. A., Imteaz, M., Choudhury, T. A. & Phillips, D. Applicability of artificial neural network in hydraulic experiments using a new sewer overflow screening device. *Australas. J. Water Resour.* **17**, 77–86 (2013).
21. Ghorbani, M. A., Jabehdar, M. A., Yaseen, Z. M. & Inyurt, S. *Solving the Pan Evaporation Process Complexity Using the Development of Multiple Mode of Neurocomputing Models* (2021).
22. Sudheer, K. P., Gosain, A. K., Mohana Rangan, D. & Saheb, S. M. Modelling evaporation using an artificial neural network algorithm. *Hydrol. Process.* **16**, 3189–3202 (2002).
23. Abed, M. M., El-Shafie, A. & Osman, S. A. B. Creep predicting model in masonry structure utilizing dynamic neural network. *J. Comput. Sci.* **6**, 597 (2010).
24. Ashrafzadeh, A., Ghorbani, M. A., Biazar, S. M. & Yaseen, Z. M. Evaporation process modelling over northern Iran: Application of an integrative data-intelligence model with the krill herd optimization algorithm. *Hydrol. Sci. J.* **64**, 1843–1856 (2019).
25. Castellano-Méndez, M., González-Manteiga, W., Febrero-Bande, M., Prada-Sánchez, J. M. & Lozano-Calderón, R. Modelling of the monthly and daily behaviour of the runoff of the Xallas river using Box-Jenkins and neural networks methods. *J. Hydrol.* **296**, 38–58 (2004).
26. Ashrafzadeh, A., Malik, A., Jothiprakash, V., Ghorbani, M. A. & Biazar, S. M. Estimation of daily pan evaporation using neural networks and meta-heuristic approaches. *ISH J. Hydraul. Eng.* **26**, 421–429 (2020).
27. Malik, A. *et al.* Modeling monthly pan evaporation process over the Indian central Himalayas: Application of multiple learning artificial intelligence model. *Eng. Appl. Comput. Fluid Mech.* **14**, 323–338 (2020).
28. Abudu, S., Cui, C., King, J. P., Moreno, J. & Bawazir, A. S. Modeling of daily pan evaporation using partial least squares regression. *Sci. China Technol. Sci.* **54**, 163–174 (2011).
29. Tabari, H., Marofi, S. & Sabziparvar, A.-A. Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression. *Irrig. Sci.* **28**, 399–406 (2010).
30. Keskin, M. E. & Terzi, Ö. Artificial neural network models of daily pan evaporation. *J. Hydrol. Eng.* **11**, 65–70 (2006).
31. Kişi, Ö. Evolutionary neural networks for monthly pan evaporation modeling. *J. Hydrol.* **498**, 36–45 (2013).
32. Deo, R. C., Samui, P. & Kim, D. Estimation of monthly evaporative loss using relevance vector machine, extreme learning machine and multivariate adaptive regression spline models. *Stoch. Environ. Res. Risk Assess.* **30**, 1769–1784 (2016).
33. Falamarzi, Y., Palizdan, N., Huang, Y. F. & Lee, T. S. Estimating evapotranspiration from temperature and wind speed data using artificial and wavelet neural networks (WNNs). *Agric. Water Manage.* **140**, 26–36 (2014).
34. Wang, L., Kisi, O., Zounemat-Kermani, M. & Li, H. Pan evaporation modeling using six different heuristic computing methods in different climates of China. *J. Hydrol.* **544**, 407–427 (2017).
35. Malik, A., Kumar, A. & Kisi, O. Monthly pan-evaporation estimation in Indian central Himalayas using different heuristic approaches and climate based models. *Comput. Electron. Agric.* **143**, 302–313 (2017).
36. Tezel, G. & Buyukyildiz, M. Monthly evaporation forecasting using artificial neural networks and support vector machines. *Theor. Appl. Climatol.* **124**, 69–80 (2016).
37. Alipour, A., Yarahmadi, J. & Mahdavi, M. Comparative study of M5 model tree and artificial neural network in estimating reference evapotranspiration using MODIS products. *J. Climatol.* **2014**, 1–11 (2014).
38. Hassan, M. A., Khalil, A., Kaseb, S. & Kassem, M. A. Exploring the potential of tree-based ensemble methods in solar radiation modeling. *Appl. Energy* **203**, 897–916 (2017).
39. Fan, J. *et al.* 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 (2018).
40. Francke, T., López-Tarazón, J. A. & Schröder, B. Estimation of suspended sediment concentration and yield using linear models, random forests and quantile regression forests. *Hydrol. Process. Int. J.* **22**, 4892–4904 (2008).
41. Feng, Y., Cui, N., Gong, D., Zhang, Q. & Zhao, L. Evaluation of random forests and generalized regression neural networks for daily reference evapotranspiration modelling. *Agric. Water Manage.* **193**, 163–173 (2017).
42. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
43. Zhao, Z., Chen, W., Wu, X., Chen, P. C. Y. & Liu, J. LSTM network: A deep learning approach for short-term traffic forecast. *IET Intell. Transp. Syst.* **11**, 68–75 (2017).
44. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
45. Wang, Z., Yan, W. & Oates, T. Time series classification from scratch with deep neural networks: A strong baseline. In *2017 International Joint Conference on Neural Networks (IJCNN)*, 1578–1585 (IEEE, 2017).
46. Lee, H., Pham, P., Largman, Y. & Ng, A. Unsupervised feature learning for audio classification using convolutional deep belief networks. *Adv. Neural Inf. Process. Syst.* **22**, 1096–1104 (2009).
47. Gao, Y., Hendricks, L. A., Kuchenbecker, K. J. & Darrell, T. Deep learning for tactile understanding from visual and haptic data. In *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 536–543 (IEEE, 2016).
48. Liu, J. N. K., Hu, Y., You, J. J. & Chan, P. W. Deep neural network based feature representation for weather forecasting. In *Proc. International Conference on Artificial Intelligence (ICAI)*, 1 (The Steering Committee of The World Congress in Computer Science, Computer, 2014).

49. Koprinska, I., Wu, D. & Wang, Z. Convolutional neural networks for energy time series forecasting. In *2018 International Joint Conference on Neural Networks (IJCNN)*, 1–8 (IEEE, 2018).
50. Biazar, S. M., Ghorbani, M. A. L. I. & Shahedi, K. Uncertainty of artificial neural networks for daily evaporation prediction (case study: Rasht and Manjil Stations). *J. Watershed Manage. Res.* **10**, 1–12 (2019).
51. Shaaban, A. J. & Low, K. S. Droughts in Malaysia: A look at its characteristics, impacts, related policies and management strategies. In *Water and Drainage 2003 Conference*, 28–29 (2003).
52. Yaseen, Z. M., El-Shafie, A., Jaafar, O., Afan, H. A. & Sayl, K. N. Artificial intelligence based models for stream-flow forecasting: 2000–2015. *J. Hydrol.* **530**, 829–844 (2015).
53. Abed, M., Imteaz, M. A., Ahmed, A. N. & Huang, Y. F. Application of long short-term memory neural network technique for predicting monthly pan evaporation. *Sci. Rep.* **11**, 1–19 (2021).
54. Freedman, D., Pisani, R., Purves, R. & Adhikari, A. *Statistics* (2007).
55. Hauke, J. & Kossowski, T. *Comparison of Values of Pearson's and Spearman's Correlation Coefficient on the Same Sets of Data* (2011).
56. Dalkılıç, Y., Okkan, U. & Baykan, N. Comparison of different ANN approaches in daily pan evaporation prediction. *J. Water Resour. Prot.* **6**, 319 (2014).
57. Wang, L., Kisi, O., Zounemat-Kermani, M. & Gan, Y. Comparison of six different soft computing methods in modeling evaporation in different climates. *Hydrol. Earth Syst. Sci. Discuss.* <https://doi.org/10.5194/hess-2016-247> (2016).
58. Zakaria, M. N. A., Malek, M. A., Zolkepli, M. & Ahmed, A. N. Application of artificial intelligence algorithms for hourly river level forecast: A case study of Muda River, Malaysia. *Alexand. Eng. J.* **60**, 4015–4028 (2021).
59. Rosenberry, D. O., Stannard, D. I., Winter, T. C. & Martinez, M. L. Comparison of 13 equations for determining evapotranspiration from a prairie wetland, Cottonwood Lake area, North Dakota, USA. *Wetlands* **24**, 483–497 (2004).
60. Stephens, J. C. & Stewart, E. H. A comparison of procedures for computing evaporation and evapotranspiration. *Publication* **62**, 123–133 (1963).
61. Thornthwaite, C. W. An approach toward a rational classification of climate. *Geogr. Rev.* **38**, 55–94 (1948).
62. Liaw, A. & Wiener, M. Classification and regression by Random Forest. *R News* **2**, 18–22 (2002).
63. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).
64. Peters, J. *et al.* Random forests as a tool for ecohydrological distribution modelling. *Ecol. Modell.* **207**, 304–318 (2007).
65. Wang, Z. *et al.* Flood hazard risk assessment model based on random forest. *J. Hydrol.* **527**, 1130–1141 (2015).
66. Quedraogo, I., Defourny, P. & Vanclooster, M. Application of random forest regression and comparison of its performance to multiple linear regression in modeling groundwater nitrate concentration at the African continent scale. *Hydrogeol. J.* **27**, 1081–1098 (2019).
67. Nair, V. & Hinton, G. E. Rectified linear units improve restricted boltzmann machines. In *Icml* (2010).
68. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
69. Gu, J. *et al.* Recent advances in convolutional neural networks. *Pattern Recognit.* **77**, 354–377 (2018).
70. Volpi, M. & Tuia, D. Dense semantic labeling of subdecimeter resolution images with convolutional neural networks. *IEEE Trans. Geosci. Remote Sens.* **55**, 881–893 (2016).
71. Springenberg, J. T., Dosovitskiy, A., Brox, T. & Riedmiller, M. Striving for simplicity: The all convolutional net. Preprint at <http://arXiv.org/1412.6806> (2014).
72. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. Preprint at <http://arXiv.org/1412.6980> (2014).
73. Taylor, K. E. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res. Atmos.* **106**, 7183–7192 (2001).
74. Change, I. C. The physical science basis. In *Contrib. Work. Gr. I to Fourth Assess. Rep. Intergov. Panel Clim. Change*, Vol. 996 (2007).
75. Majhi, B., Naidu, D., Mishra, A. P. & Satapathy, S. C. Improved prediction of daily pan evaporation using deep-LSTM model. *Neural Comput. Appl.* **32**, 7823–7838 (2020).
76. Abed, M., Imteaz, M. & Ali Najah Ahmed, Y. F. H. Improved prediction of monthly pan evaporation utilising support vector machine technique. In *The Asia-Pacific Conference on Computer Science and Data Engineering 2021* (IEEE).

## Acknowledgements

The author would like to thank the Australian Government Research Training Program Scholarship (RTP) for its support. Also, the Malaysian Meteorological Department (MMD) for providing this study with the data.

## Author contributions

M.A.: methodology, formal analysis, visualization and writing—review and editing, M.A.I.: writing—review and editing and supervision, A.N.A.: writing—review and editing and supervision, Y.F.H.: data curation and writing—review and editing.

## Funding

The APC was covered by Universiti Tunku Abdul Rahman (UTAR), Malaysia, via Research Publication Scheme (Project Number: UTARRPS 6251/H03) and UTARFSPJP.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Y.F.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022