Kernel Extreme Learning Machine: An Efficient Model for Estimating Daily Dew Point Temperature Using Weather Data

Meysam Alizamir 1,2,*, Sungwon Kim 3, Mohammad Zounemat-Kermani 4, Salim Heddam 5, Nam Won Kim 6, and Vijay P. Singh 7,8

1 Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
2 The Faculty of Civil Engineering, Duy Tan University, Da Nang 550000, Vietnam
3 Department of Railroad Construction and Safety Engineering, Dongyang University, Yeongju 36040, Korea; swkim1968@dyu.ac.kr
4 Department of Water Engineering, Shahid Bahonar University of Kerman, Kerman 7616913439, Iran; zounemat@uk.ac.ir
5 Faculty of Science, Agronomy Department, Hydraulics Division, Laboratory of Research in Biodiversity Interaction Ecosystem and Biotechnology, University 20 Août 1955, Route El Hadaïk, BP 26, Skikda 21000, Algeria; heddamsalim@yahoo.fr
6 Department of Land, Water and Environment Research Institute, Korea Institute of Civil Engineering and Building Technology, Goyang 10223, Korea
7 Department of Biological and Agricultural Engineering & Zachry Department of Civil & Environmental Engineering, Texas A&M University, 321 Scoates Hall, 2117 TAMU, College Station, TX 77843-2117, USA; vsingh@tamu.edu
8 National Water Center, UAE University, Al Ain 17666, UAE
* Correspondence: meysamalizamir@duytan.edu.vn (M.A.); nwkim@kict.re.kr (N.W.K.)

Received: 28 July 2020; Accepted: 13 September 2020; Published: 17 September 2020

Abstract: Accurate estimation of dew point temperature ($T_{dew}$) has a crucial role in sustainable water resource management. This study investigates kernel extreme learning machine (KELM), boosted regression tree (BRT), radial basis function neural network (RBFNN), multilayer perceptron neural network (MLPNN), and multivariate adaptive regression spline (MARS) models for daily dew point temperature estimation at Durham and UC Riverside stations in the United States. Daily time scale measured hydrometeorological data, including wind speed (WS), maximum air temperature (TMAX), minimum air temperature (TMIN), maximum relative humidity (RHMAX), minimum relative humidity (RHMIN), vapor pressure (VP), soil temperature (ST), solar radiation (SR), and dew point temperature ($T_{dew}$) were utilized to investigate the applied predictive models. Results of the KELM model were compared with other models using eight different input combinations with respect to root mean square error (RMSE), coefficient of determination ($R^2$), and Nash–Sutcliffe efficiency (NSE) statistical indices. Results showed that the KELM models, using three input parameters, VP, TMAX, and RHMIN, with $RMSE = 0.419 \degree C$, $NSE = 0.995$, and $R^2 = 0.995$ at Durham station, and seven input parameters, VP, ST, RHMAX, TMIN, RHMIN, TMAX, and WS, with $RMSE = 0.485 \degree C$, $NSE = 0.994$, and $R^2 = 0.994$ at UC Riverside station, exhibited better performance in the modeling of daily $T_{dew}$. Finally, it was concluded from a comparison of the results that out of the five models applied, the KELM model was found to be the most robust by improving the performance of BRT, RBFNN, MLPNN, and MARS models in the testing phase at both stations.

Keywords: daily dew point temperature; kernel extreme learning machine; estimation; climatic data
1. Introduction

Dew point temperature (T_{dew}) plays a vital role in the elaboration and application of several ecological, hydrological, and meteorological models, especially for the quantification of evapotranspiration [1–3]. Different important climatic parameters can be affected by T_{dew}. In addition, it has been demonstrated that T_{dew} can be used as an important factor for climate change studies [4]. Recently, Ali et al. demonstrated a strong relationship between T_{dew} and extreme precipitation [5]. Bui et al. reported that T_{dew} helped significantly to understand the relationship between precipitation and air temperature, which can be used to quantify near-surface humidity [6,7]. Several authors have paid attention to the strong relationship between T_{dew} and meteorological variables [8–11]. Many studies have tried to develop models to link T_{dew} to meteorological variables using machine learning models. Dong et al. applied machine learning models for modeling dew point temperature (T_{dew}) using several meteorological variables as inputs, namely, TMAX, TMIN, TMEAN, RHMAX, RHMIN, RHMEAN, and atmospheric pressure (Pa) [12]. The authors applied and compared ten different soft computing techniques to model T_{dew}. Based on the results, the best accuracy was obtained by the Bat–ELM model by employing Tmax, Tmin, RHmax, and RHmin as input variables. Shiri presented three data-driven models to model T_{dew} at weekly and daily time scales using data collected at six meteorological stations [13]. The proposed three models were gene expression programming (GEP), MARS, and RF. For modeling T_{dew}, Tmean, RHmean, sunshine hours (SH), and wind speed (U2) were used as input parameters. In the second application, the previously measured values of T_{dew} were applied as input variables for modeling T_{dew} one day and seven days in advance. The best accuracy was obtained using the MARS model by employing Tmean, RHmean, and SH as input variables at all stations, respectively, while for T_{dew} prediction, GEP worked best at daily and weekly time steps. Qasem et al. employed three machine learning models for estimating T_{dew}, namely, M5 model tree (M5Tree), GEP, and SVM, using Tmean, RHmean, SH, U2, and actual vapor pressure (VP) [14]. A comparison between the models revealed that the M5Tree model, by having five climatic variables (Tmean, RHmean, SH, U2, and VP), had the best accuracy. Naganna et al. introduced new hybrid models for estimating daily T_{dew} using bulb temperature (TB), VP, and RHmean [15]. They applied MLPNN optimized using the gravitational search algorithm (MLPNN–GSA) and the firefly algorithm (MLPNN–FFA). Results obtained using MLPNN–FFA and MLPNN–GSA were compared to those obtained using the standard MLPNN, SVM, and ELM models, and the best accuracy was yielded by the MLPNN–FFA model.

Attar et al. compared MARS, GEP, and SVM models for predicting T_{dew} using Tmax, Tmin, RHmean, SH, U2, and atmospheric pressure (P) [16]. The best accuracy was obtained by the MARS model. Mehdizadeh et al. applied the GEP model for modeling and forecasting T_{dew} using several meteorological variables, namely, Tmax, Tmin, Tmean, VP, RHmean, and P [17]. The study was conducted according to three scenarios: (i) temperature-based models using only air temperature as an input variable, (ii) using a combination of meteorological variables, and (iii) forecasting models using previously measured values of T_{dew} as input variables. For the first scenario, the best accuracy was obtained using Tmin and the difference between Tmin and Tmax as input variables. For the second scenario, the best accuracy was obtained using VP, RHmean, and P. Finally, for the third scenario, the best accuracy was obtained by the model with T_{dew} measured on the three previous days and the Julian day (J) as input variables. In some regions of India, Deka et al. compared the SVM and ELM models for modeling T_{dew} using daily measured TB, VP, and RHmean [18]. They reported that the ELM model was more accurate than the SVM model and achieved the best accuracy. Shiri et al. introduced a new artificial neural network model called the Elman discrete recurrent neural network (EDRNN) model, trained using two different algorithms: (i) the conjugate gradient learning algorithm and (ii) the quick prop learning algorithm [19]. Results obtained using the EDRNN model were compared to those obtained using the GEP model. The three models were developed using Tmean, RHmean, U2, P, and solar radiation (SR). By comparing several input combinations, the authors demonstrated that GEP significantly surpassed the two ANN models.
Kisi et al. investigated the accuracy of four machine learning models, namely, ANFIS-C, ANFIS-G, GRNN, and SOM [20]. Another study was conducted in South Korea using Tmean, RHmean, U2, SH, and VP parameters. The authors reported two important conclusions. First, none of U2, SH, and VP contributed significantly to the improvement of the machine learning models, while Tmean and RHmean were the most efficient variables for predicting T\textsubscript{dew}. Furthermore, results revealed that the best accuracy was obtained using the GRNN model, while the lowest accuracy was achieved using the SOM model. A hybrid model by combining SVM and the firefly algorithm (SVM–FFA) was developed by Al-Shammari et al. for modeling T\textsubscript{dew} using Tmean, RHmean, and P, measured at daily time scale in Iran [21]. Results obtained using the SVM–FFA model were compared to those obtained using MLPNN, SVM, and genetic programming (GP), and the best accuracy was obtained using SVM–FFA.

Amirmojahedi et al. employed a new hybrid model by combining the wavelet transform and the extreme learning machines (W–ELM) for predicting daily T\textsubscript{dew} using Tmean, RHmean, and P [22]. Compared to the standard ELM, SVM, and MLPNN models, W-ELM yielded the best result, while the lowest accuracy was obtained using the MLPNN model. In another study, Baghban et al. applied the least square support vector machine optimized by genetic algorithm (LSSVM–GA) for modeling T\textsubscript{dew} using Tmean, RHmean, and P [23]. They found that the LSSVM–GA was more accurate compared to ANFIS–GA. In summary, although many studies have been conducted for modeling T\textsubscript{dew}, this study investigates a reliable tool, the kernel extreme learning machine (KELM), to estimate daily T\textsubscript{dew} by using hydrometeorological input parameters. Results obtained using KELM were compared to those obtained using the boosted regression tree (BRT), radial basis function neural network (RBFNN), MARS, and MLPNN models. To the best of the authors’ knowledge, this is the first study that applies the KELM model to estimate daily T\textsubscript{dew} at the Durham and UC Riverside stations in the USA.

2. Materials and Methods

2.1. Artificial Neural Networks (MLPNN and RBFNN)

Artificial neural networks (ANNs) are computational networks and information processing systems that consist of a large number of interconnected computing units called neurons. The formation of ANNs is based on the mathematical simulation of the biological nervous system structure. To date, various versions of ANNs have been developed. In this study, two common types of ANNs, including multilayer perceptron (MLPNN) and radial basis (RBNN), which have had proper performance in simulating hydrometeorological issues, are reviewed [24,25]. Both the MLPNN and RBNN models are neural networks with a supervised network training structure and feed-forward information transfer. In both models, the developed network consists of an input layer, a hidden layer, and an output layer. It is worth mentioning that although MLPNNs can be developed with more than one layer in the hidden layer, these models are often developed with only one layer [26–30]. The input layer receives user-entered information and transmits it to the hidden layer after applying the weight coefficient and biases. The neurons in the hidden layer, defined according to the activation function, process the receive values from the input layer neurons and send them to the output layer neurons. Similar to the hidden layer neurons, the neurons in the output layer compute the model output by using an activation function [2,31]. Figure 1 shows the general structure of an MLPNN with two hidden layers, as well as the general structure of an RBNN.

In this phase, the computed output values of the network are compared with the target values. The main purpose of this network is to minimize the discrepancy between computational and observational values in the mean square error format (as an error function). To this end, back-propagation-based methods, such as gradient descent algorithms, are used. The main difference between the two methods of MLPNN and RBNN can be considered in the type of activation functions.
The sigmoid functions are generally used in MLPNN networks (Equation (1)), and the radial basis Gaussian function (Equation (2)) is employed in RBNN networks.

\[ f(x) = \frac{1}{1 + \exp(-x)} \]  
\[ g(x) = \exp\left(-\frac{(x-c)^2}{\sigma^2}\right) \]

where \( x \) represents the input variable, \( c \) is the center, and \( \sigma \) is the variance. Further detailed information on the basics of developing and deploying MLPNN and RBNN networks has been provided in detail in many other references [32–34]. In this study, the Levenberg–Marquardt (LM) learning method is applied to train the network based on its fast convergence capability for complex datasets. Additionally, Sigmoid and linear transfer functions are employed for hidden and output layers, respectively. Moreover, by applying a trial and error process, the number of neurons in the hidden layer is found to be 10 and 12 for MLPNN and RBFNN models, respectively.

![Figure 1](image.png)

**Figure 1.** The general structure of (a) multilayer perceptron and (b) radial basis function neural network.

### 2.2. Kernel Extreme Learning Machine (KELM)

Even though the artificial neural network models have proven their high capabilities in modeling plenty of nonlinear engineering problems, such as dew point temperature, they might suffer from the drawback of using gradient-descent-based training algorithms. That is, these algorithms can be trapped in local minima. Moreover, the ANNs’ architecture contains many processors (neurons) as well as network parameters, which make it a sophisticated black-box structure in comparison to other machine learning models. To cope with these weaknesses of regular ANN models, Huang et al. proposed a novel version of a training algorithm called extreme learning machine (ELM) [35]. ELMs are single hidden feed-forward layer (SLFN) networks that choose the input weights and biases randomly, whereas the output hidden neurons are calculated by the Moore–Penrose generalized inverse. The ELM models have been successfully used in several applications in hydrological modeling in the past couple of years [36,37]. Nonetheless, the standard version of the ELM model faces the drawback of providing different accuracies in various trials due to its randomly assigned weight strategy. To resolve this shortcoming of the standard ELM, Huang et al. proposed kernel ELM (KELM) by modifying its random process of allocating random weights between the input and hidden layers [38]. This section briefly explains the theory and methodology of KELM. Complete details for the KELM model can be found in Huang et al.’s paper [38]. To begin with, one can present the formulation of the general
structure of an SLFN model having \( M \) training samples, \( N \) hidden nodes, and \( g(x) \) as the activation function, as the following \([38,39]\):

\[
\sum_{i=1}^{N} \beta_i g(w_i x_j + b_i) = o_k; \quad i = 1, 2, \ldots, M
\]  

(3)

In Equation (3), \( i \) corresponds to a hidden node, and \( j \) denotes each independent input variable. \( o \) is the output vector, \( x \) represents the input feature vector, and \( b \) is the bias. \( w \) shows the weight vector between the input layer (IL) and the hidden layer (HL), while \( \beta \) denotes the weight vector for connecting the nodes in the hidden layer to output nodes. Assuming an ideal condition for the developed SLFN model in Equation (1), one can expect zero errors between the target value (\( t \)) and the SLFN model’s output (\( o \)). In that case, Equation (3) can be rewritten as below:

\[
\sum_{i=1}^{N} \beta_i g(w_i x_j + b_i) = t_k
\]  

(4)

Consequently, it is possible to arrange Equation (4) in the form of \( \mathbf{H}\beta = \mathbf{T} \). \( \mathbf{H} \) and \( \mathbf{T} \) are the hidden layer output matrix and the activation function matrix, respectively. Equation (2) can be solved using linear methods, for example, the MP-generalized inverse of \( \mathbf{H} \), known as \( \mathbf{H}^\dagger \).

\[
\mathbf{H}\beta = \mathbf{T} \Rightarrow \hat{\beta} = \mathbf{H}^\dagger \mathbf{T}
\]  

(5)

In the KELM model, a kernel, such as \( K(x, y) \), maps the data from IL to the HL space. In this sense, the KELM model applies the orthogonal projection procedure to compute the \( \mathbf{H}^\dagger \) matrix so that it can be written as

\[
K_{ELM_{ij}} = \mathbf{H}\mathbf{H}^\dagger : K(x_i, x_j) = h(x_i)h(x_j)
\]  

(6)

For this purpose, the reliable Gaussian kernel is applied for mapping the data between the layers.

\[
K(x, x_i) = \exp\left(-\gamma \|x - x_i\|^2\right)
\]  

(7)

where \( \gamma \) is the kernel parameter. In this study, the regularization coefficient and type of kernel are set to 35 and wavelet kernel, respectively.

The step procedures of the developed KELM model in this study is as follows:

- Dividing the dataset into the training (80%) and testing (20%) sets.
- Constructing an SLFN model on the training set (\( M \) training samples).
- Assigning random values for the input weights and biases (\( w_i \) and \( b_i \)).
- Calculating the HL output matrix (\( \mathbf{H} \)).
- Calculating \( \beta \) as the output weight.
- Adopting the kernel matrix for ELM (\( K_{ELM} \)).
- Calculate the output function of KELM based on the Gaussian radial basis function.
- Evaluating the developed KELM performance for the testing data set.

2.3. Multivariate Adaptive Regression Splines (MARS)

MARS is a learning machine method consisting of several simple regression models that have high capability in estimating and simulating complex phenomena. In this method, the space issue is divided into intervals of predictive variables. In each interval, the spline functions are fitted to the existing data. The formation of a MARS model is based on the creation of piecewise linear basis functions.
\[ |x - t|_+ = \max(0, x - t) = \begin{cases} x - t, & x > t \\ 0, & x \leq t \end{cases} \]
\[ |t - x|_+ = \max(0, t - x) = \begin{cases} t - x, & x < t \\ 0, & x \geq t \end{cases} \]  

(8)

where \( t \) represents the knot in the MARS model. Given the written form of Equations (8), the equations written above are known as reflected-pair functions [40]. Therefore, having the \( n \) input variable \( X \), BFs can be expressed according to the following equations:

\[ C = \left[ (X_j - t)_{+}, (t - X_j)_{+} \right] : t \in \{ x_1, x_2, \ldots, x_n \} \]  

(9)

where \( n \) is the total number of observations, and \( j = 1, 2, \ldots, p \). The basis functions (BFs) in the MARS method include the input variable, and they express the relationship between the input variables and the target parameter. Finally, by combining the generated spline functions, a resilient and efficient model is created to predict the target parameter as follows:

\[ y = \beta_0 + \sum_{i=1}^{M} \beta_i B_i(X) \]  

(10)

In the above equation, \( B_i(X) \) are BFs in \( C \) or the product of two or more \( C \) functions (see Equation (9)). \( \beta_0, \beta_i \) represent the bias and the coefficients of BFs, which can be calculated by the least square method (LSM), and \( M \) stands for the number of terms in a forward/backward stepwise process [29,41,42].

2.4. Boosted Regression Tree (BRT)

BRTs are tree models capable of simulating complex processes, which are formed by combining statistical methods and machine learning. The BRT method is based on the creation of an ensemble learning structure called boosting. For this purpose, the performance of a number of developed premise tree models (such as classification and regression trees (CART)) is aggregated by using a boosting technique, which results in better performance of the BRT model than the single CART model in the simulation and prediction of the desired phenomenon [43]. In the boosting method used in BRT, a forward stepwise process is applied so that each of the CART models formed is only assigned to a subset of training data. The subsets are selected by a stochastic process without the possibility of replacement. In the BRT model, two key parameters, including the learning rate and tree complexity factor, are responsible for creating and forming the overall structure of the model. The learning rate parameter controls the number of nodes in each tree to determine the contribution of each CART to the overall BRT structure and tree complexity factor [44].

3. Description of Study Area and Observational Data

In this research, daily hydroclimatic parameters, including wind speed (WS), maximum air temperature (TMAX), minimum air temperature (TMIN), maximum relative humidity (RHMAX), minimum relative humidity (RHMIN), vapor pressure (VP), soil temperature (ST), and solar radiation (SR) were applied for estimating the values of daily dew point temperature (\( T_{dew} \)) at both stations: Durham (latitude 39°36′ N, longitude 121°49′ W, and altitude 39.6 m) is located in Sacramento Valley Region Butte County, and UC Riverside (latitude 33°57′ N, longitude 117°20′ W, and altitude 310.9 m) is in the Los Angeles Basin Region Riverside County. The locations of Durham and UC Riverside Stations can be found in Figure 2. The data used in this study were obtained from the California Irrigation Management Information System (CIMIS). CIMIS weather station dataset quality was controlled by applying different data processing, including analyzing the measured weather data accuracy (https://cimis.water.ca.gov/cimis/). The data from 1 January 2005 to 31 December 2009 were applied in this study. In the study, the training dataset included the first 80% of daily data, and the
residual data were applied for the testing of the models. Table 1 presents statistical parameters such as average, minimum (Min), maximum (Max), and standard deviation (St. Dev) of the weather variables employed in this study.

![Figure 2](imageurl). The locations of the Durham and UC Riverside stations in the state of California, USA.

**Table 1.** Basic statistical properties of the used dataset.

| Station     | Data Set   | Input and Output Parameters | Unit | Average Min. | Max. | St. Dev. |
|-------------|------------|------------------------------|------|--------------|------|----------|
| Durham      | Training   | Wind speed                  | m/s  | 1.79         | 0.50 | 7.50     | 0.87     |
|             |            | Max. temperature            | °C   | 23.10        | 4.00 | 40.00    | 8.14     |
|             |            | Min. temperature            | °C   | 8.41         | −7.00| 25.70    | 5.76     |
|             |            | Max. relative humidity      | %    | 88.82        | 29.00| 100.00   | 9.33     |
|             |            | Min. relative humidity      | %    | 40.31        | 0    | 98.00    | 19.23    |
|             |            | Vapor pressure              | Pa   | 1.16         | 0.20 | 2.90     | 0.46     |
|             |            | Soil temperature            | °C   | 16.66        | 3.60 | 29.60    | 6.47     |
|             |            | Solar radiation             | MJ/m²| 197.48       | 2.00 | 460.00   | 100.52   |
|             | Testing    | Wind speed                  | m/s  | 1.71         | 0.60 | 5.20     | 0.85     |
|             |            | Max. temperature            | °C   | 22.82        | 4.90 | 38.80    | 8.00     |
|             |            | Min. temperature            | °C   | 8.08         | −7.80| 20.40    | 5.80     |
|             |            | Max. relative humidity      | %    | 91.95        | 43.00| 100.00   | 8.98     |
|             |            | Min. relative humidity      | %    | 42.51        | 4.00 | 97.00    | 18.43    |
|             |            | Vapor pressure              | Pa   | 1.20         | 0.40 | 2.30     | 0.47     |
|             |            | Soil temperature            | °C   | 16.80        | 4.20 | 27.70    | 6.75     |
|             |            | Solar radiation             | MJ/m²| 196.09       | 10.00| 385.00   | 98.02    |
|             |            | Dew point temperature       | °C   | 8.19         | −17.00| 23.90    | 6.03     |
| UC Riverside| Training   | Wind speed                  | m/s  | 1.75         | 0.40 | 6.20     | 0.59     |
|             |            | Max. temperature            | °C   | 25.48        | 7.70 | 43.80    | 7.17     |
|             |            | Min. temperature            | °C   | 11.52        | −3.20| 25.70    | 4.85     |
|             |            | Max. relative humidity      | %    | 72.14        | 15.00| 91.00    | 16.69    |
|             |            | Min. relative humidity      | %    | 29.44        | 4.00 | 83.00    | 15.62    |
|             |            | Vapor pressure              | Pa   | 1.01         | 0.10 | 2.20     | 0.39     |
|             |            | Soil temperature            | °C   | 17.32        | 5.50 | 27.50    | 5.09     |
|             |            | Solar radiation             | MJ/m²| 205.80       | 0.00 | 428.00   | 79.14    |
|             |            | Dew point temperature       | °C   | 6.02         | −18.70| 18.80    | 6.73     |
|             | Testing    | Wind speed                  | m/s  | 1.81         | 0.90 | 4.90     | 0.52     |
|             |            | Max. temperature            | °C   | 25.35        | 9.80 | 40.60    | 7.11     |
|             |            | Min. temperature            | °C   | 11.61        | 0.30 | 22.10    | 4.69     |
|             |            | Max. relative humidity      | %    | 69.32        | 18.00| 86.00    | 16.95    |
|             |            | Min. relative humidity      | %    | 28.50        | 4.00 | 81.00    | 15.21    |
|             |            | Vapor pressure              | Pa   | 0.97         | 0.20 | 1.80     | 0.38     |
|             |            | Soil temperature            | °C   | 17.47        | 7.40 | 26.10    | 5.20     |
|             |            | Solar radiation             | MJ/m²| 199.66       | 17.00| 313.00   | 71.00    |
|             |            | Dew point temperature       | °C   | 5.39         | −13.10| 15.80    | 6.64     |
4. Performance Indices

For the reliability assessment of models, three evaluation indicators, root mean squared error (RMSE), coefficient of determination ($R^2$), and the Nash–Sutcliffe efficiency coefficient (NSE), were applied to assess the performance of the five models developed for estimating dew point temperature ($T_{dew}$). These criteria are defined as

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} ((T_{dew})_{io} - (T_{dew})_{ip})^2}{n}}; \text{ Range } = [0, \infty], \text{ Ideal value } = 0 \tag{11}
\]

\[
R^2 = \frac{\sum_{i=1}^{n} ((T_{dew})_{io} - (T_{dew})_{ip})^2}{\sum_{i=1}^{n} ((T_{dew})_{io} - (T_{dew})_{io})^2 \sum_{i=1}^{n} ((T_{dew})_{ip} - (T_{dew})_{ip})^2}; \text{ Range } = [0, 1], \text{ Ideal value } = 1 \tag{12}
\]

\[
NS = 1 - \frac{\sum_{i=1}^{n} ((T_{dew})_{io} - (T_{dew})_{ip})^2}{\sum_{i=1}^{n} ((T_{dew})_{io} - (T_{dew})_{io})^2}; \text{ Range } = [-\infty, 1], \text{ Ideal value } = 1 \tag{13}
\]

where $n$ represents the number of data, $(T_{dew})_{io}$ stands for the observed dew point temperature values, and $(T_{dew})_{ip}$ is the model’s estimates.

5. Results and Discussion

Table 1 expresses the descriptive statistics of the observed dataset, including training and testing data from Durham and UC Riverside stations, respectively. The values of the standard deviation criterion produced considerable fluctuations within the used datasets. The results of the correlation matrix between dew point temperature and input parameters are provided in Table 2 for Durham and UC Riverside stations, respectively. Vapor pressure (VP) was the best parameter for correlating dew point temperature ($T_{dew}$), whereas wind speed (WS) was identified as the inverse parameter for correlating dew point temperature ($T_{dew}$) at both stations.

Table 2. The correlation matrix between dew point temperature and input parameters.

| Correlation          | Stations         | Output | Inputs                  |  |  |  |  |  |  |  |  |  |  |  |  |
|----------------------|------------------|--------|-------------------------|---|---|---|---|---|---|---|---|---|---|---|---|
|                      |                  |        | WS | Max. Temp. | Min. Temp. | Max. RH | Min. RH | VP | ST | SR |  |  |  |  |
| Pearson’s correlation| Durham           | $T_{dew}$ | -0.294 | 0.693 | 0.833 | 0.261 | 0.183 | 0.971 | 0.792 | 0.492 |
| Spearman’s correlation|                 |         | -0.271 | 0.719 | 0.86 | 0.02 | 0.203 | 0.997 | 0.814 | 0.525 |
| Pearson’s correlation| UC Riverside     | $T_{dew}$ | -0.163 | 0.387 | 0.659 | 0.675 | 0.487 | 0.968 | 0.758 | 0.382 |
| Spearman’s correlation|                 |         | -0.271 | 0.719 | 0.86 | 0.02 | 0.203 | 0.997 | 0.814 | 0.525 |

5.1. Durham Station

Based on climatic parameters for predicting dew point temperature ($T_{dew}$), different parameters, including wind speed (WS), maximum temperature (TMAX), minimum temperature (TMIN), maximum relative humidity (RHMAX), minimum relative humidity (RHMIN), vapor pressure (VP), soil temperature (SR), and solar radiation (SR), were chosen. The input parameters were framed as one-, two-, three-, four-, five-, six-, seven-, and eight-input composites. The performance metrics (i.e., RMSE, NSE, and $R^2$) for the best models are shown in Table 3 for Durham station. In one-input composite
models, the values of three evaluation indicators for the vapor pressure (VP) parameter were the best in all developed models. It is clear from Table 3 that the KELM model \( \text{RMSE} = 0.426 \degree C, \text{NSE} = 0.995, \) and \( R^2 = 0.995 \) slightly surpassed the BRT, MARS, RBFNN, and MLPNN models during the testing phase. Therefore, vapor pressure (VP) among different parameters was chosen as the major parameter to formulate two-, three-, four-, five-, six-, seven-, and eight-input composite models. In two-input composite models, the composite of VP and TMAX parameters was better than the other two-input composites in the BRT, MARS, RBFNN, and MLPNN models except for the KELM model (i.e., the composite of VP and SR). It can be suggested from Table 3 that the BRT and RBFNN models slightly outperformed the MARS and MLPNN models during the testing phase. Additionally, the KELM model \( \text{RMSE} = 0.426 \degree C, \text{NSE} = 0.995, \) and \( R^2 = 0.995 \) had the best performance among other two-input composite models during the testing phase. Among three-input composite models, the composite of VP, TMAX, and RHMIN parameters was better than the other three-input composites in the BRT, RBFNN, MLPNN, and KELM models except for the MARS model (i.e., the composite of VP, ST, and RHMAX parameters). The KELM model \( \text{RMSE} = 0.419 \degree C, \text{NSE} = 0.995, \) and \( R^2 = 0.995 \) yielded the best performance among other three-input composite models during the testing phase. By comparing four-input composite models, it was found that the composite of VP, TMIN, ST, and RHMIN parameters was better than the other four-input composites in the MARS, RBFNN, and KELM models, whereas the composite of VP, TMIN, ST, and RHMAX parameters was the best in the BRT and MLPNN models. It can be seen that the MLPNN and KELM models exceeded the BRT, MARS, and RBFNN models during the testing phase. Additionally, the KELM model \( \text{RMSE} = 0.435 \degree C, \text{NSE} = 0.995, \) and \( R^2 = 0.995 \) gave, confidently, the best performance among the other four-input composite models during the testing phase.

Table 3. Results of the best BRT, MARS, RBFNN, MLPNN, and KELM models for estimating daily \( T_{dew} \) at Durham Station.

| Input Combination    | Model | Testing Phase | RMSE (°C) | NSE | \( R^2 \) |
|----------------------|-------|---------------|-----------|-----|----------|
| One-input            | VP    | BRT           | 0.460     | 0.994 | 0.994    |
|                      |       | MARS          | 0.526     | 0.992 | 0.992    |
|                      |       | RBFNN         | 0.433     | 0.995 | 0.995    |
|                      |       | MLPNN         | 0.428     | 0.995 | 0.995    |
|                      |       | KELM          | 0.426     | 0.995 | 0.995    |
| Two-input            | VP, TMAX | BRT       | 0.458     | 0.994 | 0.994    |
|                      |       | MARS          | 0.579     | 0.991 | 0.991    |
|                      |       | RBFNN         | 0.442     | 0.994 | 0.995    |
|                      |       | MLPNN         | 0.481     | 0.994 | 0.994    |
|                      |       | KELM          | 0.426     | 0.995 | 0.995    |
| Three-input          | VP, TMAX, RHMIN  | BRT       | 0.463     | 0.994 | 0.994    |
|                      |       | MARS          | 0.524     | 0.992 | 0.993    |
|                      |       | RBFNN         | 0.456     | 0.994 | 0.994    |
|                      |       | MLPNN         | 0.427     | 0.995 | 0.995    |
|                      |       | KELM          | 0.419     | 0.995 | 0.995    |
| Four-input           | VP, TMIN, ST, RHMAX  | BRT       | 0.505     | 0.993 | 0.993    |
|                      |       | MARS          | 0.541     | 0.992 | 0.992    |
|                      |       | RBFNN         | 0.597     | 0.990 | 0.990    |
|                      |       | MLPNN         | 0.494     | 0.993 | 0.993    |
|                      |       | KELM          | 0.435     | 0.995 | 0.995    |
Table 3. Cont.

| Input Combination       | Model | RMSE (°C) | NSE  | $R^2$ |
|-------------------------|-------|-----------|------|-------|
| **Five-input**          |       |           |      |       |
| VP, TMIN, ST, TMAX, RHMAX | BRT   | 0.509     | 0.993| 0.993 |
| VP, TMIN, ST, TMAX, RHMAX | MARS  | 0.528     | 0.992| 0.992 |
| VP, TMIN, ST, TMAX, RHMAX | RBFNN | 0.546     | 0.992| 0.992 |
| VP, TMIN, ST, TMAX, RHMAX | MLPNN | 0.463     | 0.994| 0.994 |
| VP, TMIN, ST, TMAX, RHMIN | KELM  | 0.423     | 0.995| 0.995 |
| **Six-input**           |       |           |      |       |
| VP, TMIN, ST, TMAX, SR, WS | BRT   | 0.508     | 0.993| 0.993 |
| VP, TMIN, ST, TMAX, SR, RHMIN | MARS  | 0.515     | 0.993| 0.993 |
| VP, TMIN, ST, TMAX, SR, RHMIN | RBFNN | 0.767     | 0.984| 0.984 |
| VP, TMIN, ST, TMAX, SR, RHMIN | MLPNN | 0.493     | 0.993| 0.993 |
| VP, TMIN, ST, TMAX, SR, RHMIN | KELM  | 0.426     | 0.995| 0.995 |
| **Seven-input**         |       |           |      |       |
| VP, TMIN, ST, TMAX, SR, WS, RHMIN | BRT   | 0.518     | 0.993| 0.993 |
| VP, TMIN, ST, TMAX, SR, WS, RHMIN | MARS  | 0.466     | 0.994| 0.994 |
| VP, TMIN, ST, TMAX, SR, WS, RHMIN | RBFNN | 0.695     | 0.987| 0.987 |
| VP, TMIN, ST, TMAX, SR, WS, RHMIN | MLPNN | 0.497     | 0.993| 0.993 |
| VP, TMIN, ST, TMAX, SR, WS, RHMIN | KELM  | 0.426     | 0.995| 0.995 |
| **Eight-input**         |       |           |      |       |
| VP, TMIN, ST, TMAX, SR, WS, RHMAX, RHMIN | BRT   | 0.632     | 0.989| 0.989 |
| VP, TMIN, ST, TMAX, SR, WS, RHMAX, RHMIN | MARS  | 0.437     | 0.995| 0.995 |
| VP, TMIN, ST, TMAX, SR, WS, RHMAX, RHMIN | RBFNN | 1.691     | 0.925| 0.927 |
| VP, TMIN, ST, TMAX, SR, WS, RHMAX, RHMIN | MLPNN | 0.595     | 0.990| 0.991 |
| VP, TMIN, ST, TMAX, SR, WS, RHMAX, RHMIN | KELM  | 0.429     | 0.995| 0.995 |

The best models yielding the lowest RMSE are shown in boldface.

Five-input composite models showed that the composite of VP, TMIN, ST, TMAX, and RHMAX parameters was better than the other five-input composites in the BRT, MARS, RBFNN, and MLPNN models except for the KELM model (i.e., the composite of VP, TMIN, ST, TMAX, and RHMIN parameters). In addition, the KELM model ($RMSE = 0.423 \, ^\circ C$, $NSE = 0.995$, and $R^2 = 0.995$) obviously achieved the best performance compared to other five-input composite models during the testing phase.

The performance of six-input composite models, with the composite of VP, TMIN, ST, TMAX, SR, and RHMIN parameters, was better than that with the other six-input composites in the MARS, RBFNN, MLPNN, and KELM models, except for the BRT model (i.e., the composite of VP, TMIN, ST, TMAX, SR, and WS parameters). Additionally, the KELM model ($RMSE = 0.426 \, ^\circ C$, $NSE = 0.995$, and $R^2 = 0.995$) attained the best performance among the other six-input composite models during the testing phase. In the case of seven- and eight-input composite models, the predictive results of three evaluation indicators revealed that the composite of VP, TMIN, ST, TMAX, SR, WS, and RHMIN parameters was better than the other seven-input composites in the BRT, RBFNN, MLPNN, and KELM models. Only the MARS model ($RMSE = 0.466 \, ^\circ C$, $NSE = 0.994$, and $R^2 = 0.994$) was accurate with the composite of VP, TMIN, ST, TMAX, SR, WS, and RHMAX parameters compared to the other seven-input composites. The KELM model ($RMSE = 0.426 \, ^\circ C$, $NSE = 0.995$, and $R^2 = 0.995$) obtained the best performance among the other seven-input composite models during the testing phase. The performance of the eight-input composite models confirmed that the KELM model ($RMSE = 0.429 \, ^\circ C$, $NSE = 0.995$, and $R^2 = 0.995$) yielded better results than the BRT, MARS, RBFNN, and MLPNN models during the testing phase. Considering the best composite models, the best performance of the developed models (i.e., BRT (two-input), MARS (eight-input), RBFNN (one-input), MLPNN, and KELM (three-input)) were found based on different composites of input parameters during the testing phase. It can be seen from Table 3 that the optimized structures of all input composites for the KELM model yielded better performance than those of BRT, MARS, RBFNN, and MLPNN models during the testing phase. Thus,
the KELM model is more influential than the BRT, MARS, RBFNN, and MLPNN models in predicting and generalizing the time series of dew point temperature at Durham station.

The scatter plots of observed and estimated daily $T_{dew}$ is shown in Figure 3 using the best composite models during the testing phase at Durham station. It can be found from the $R^2$ values that there is a slight difference among the BRT, MARS, RBFNN, MLPNN, and KELM models. The KELM model exhibited better performance than the BRT, MARS, RBFNN, and MLPNN models, while the BRT model yielded the lowest accuracy among the best composite models at Durham station.

**Figure 3.** Comparison of observed and estimated daily $T_{dew}$ using the best composite models (Durham station).

Figure 4 illustrates a comparison of RMSE values for the best composite models during the testing phase at Durham station. It can be seen from Figure 4 that the RMSE values of BRT, MARS, and RBFNN models were larger than those of MLPNN and KELM models during the testing phase. In addition, the KELM model produced the best accuracy, whereas the MARS model produced the lowest accuracy based on the best composite models at Durham station.
Figure 4. Comparison of root mean square error (RMSE) values for the best composite models during the testing phase (Durham station).

Figure 5 explains the error histogram comprising mean ($\mu$) and standard deviation ($\sigma$) for the best composite models at Durham station. A comparison indicated that the KELM model provided the lowest standard deviation, whereas the BRT model supplied the highest standard deviation based on the best composite models during the testing phase. This follows the pattern of RMSE values in the best composite models during the testing phase at Durham station.

![Error Histograms](image)

Figure 5. Error histogram of the best composite models (a) BRT, (b) MLPNN, (c) RBFNN, (d) MARS, and (e) KELM; Durham station.)
5.2. UC Riverside Station

The performance metrics of the best composite models are provided in Table 4 for UC Riverside station. For one-input composite models, the values of three evaluation indicators for the vapor pressure (VP) parameter were the best for the BRT, MARS, RBFNN, MLPNN, and KELM models during the testing phase. As seen from Table 4, the KELM model \( (RMSE = 0.570 \degree C, NSE = 0.992, \text{ and } R^2 = 0.992) \) surpassed the BRT, MARS, RBFNN, and MLPNN models during the testing phase. Therefore, vapor pressure (VP) among different parameters was chosen as the basic parameter to define two-, three-, four-, five-, six-, seven-, and eight-input composite models. In two-input composite models, the composite of VP and ST parameters was better than other two-input composites in the BRT and RBFNN models, and the composite of VP and RHMAX parameters was better than the other two-input composites in the MARS and KELM models. In addition, the MLPNN model \( (RMSE = 0.581 \degree C, NSE = 0.992, \text{ and } R^2 = 0.992) \) had an excellent performance based on the composite of VP and TMIN parameters. Additionally, the KELM model \( (RMSE = 0.573 \degree C, NSE = 0.992, \text{ and } R^2 = 0.992) \) provided the best performance among the other two-input composite models during the testing phase. Among the three-input composite models, the composite of VP, ST, and RHMAX parameters was better than the other three-input composites in the BRT, MLPNN, and KELM models. Among them, the KELM model \( (RMSE = 0.559 \degree C, NSE = 0.992, \text{ and } R^2 = 0.992) \) exceeded the BRT and MLPNN models during the testing phase. Additionally, the RBFNN model \( (RMSE = 0.572 \degree C, NSE = 0.992, \text{ and } R^2 = 0.992) \) gave outstanding accuracy for the composite of VP, ST, and RHMIN parameters. In addition, the MARS model \( (RMSE = 0.642 \degree C, NSE = 0.990, \text{ and } R^2 = 0.991) \) provided superior performance with the composite of VP, RHMAX, and SR parameters.

Table 4. Results of the best BRT, MARS, RBFNN, MLPNN, and KELM models for estimating daily \( T_{\text{dew}} \) at UC Riverside Station.

| Input Combination | Model | Testing Phase |
|-------------------|-------|---------------|
|                   |       | RMSE (°C) | NSE | \( R^2 \) |
| One-input         | BRT   | 0.592      | 0.992 | 0.992    |
|                   | MARS  | 0.671      | 0.989 | 0.989    |
|                   | VP    | 0.595      | 0.991 | 0.991    |
|                   | RBFNN | 0.595      | 0.991 | 0.991    |
|                   | MLPNN | 0.570      | 0.992 | 0.992    |
|                   | KELM  |            |       |          |
| Two-input         | VP, ST| BRT         | 0.596 | 0.991 | 0.991    |
|                   | VP, RHMAX | MARS     | 0.655 | 0.990 | 0.990    |
|                   | VP, ST | RBFNN      | 0.581 | 0.992 | 0.992    |
|                   | VP, TMIN | MLPNN  | 0.581 | 0.992 | 0.992    |
|                   | VP, RHMAX | KELM   | 0.573 | 0.992 | 0.992    |
| Three-input       | VP, ST, RHMAX | BRT     | 0.603 | 0.991 | 0.991    |
|                   | VP, RHMAX, SR | MARS  | 0.642 | 0.990 | 0.991    |
|                   | VP, ST, RHMIN | RBFNN | 0.572 | 0.992 | 0.992    |
|                   | VP, ST, RHMAX | MLPNN | 0.597 | 0.991 | 0.992    |
|                   | VP, ST, RHMAX, RHMIN | KELM | 0.559 | 0.992 | 0.992    |
| Four-input        | VP, ST, RHMAX, RHMIN, WS | BRT | 0.598 | 0.991 | 0.991    |
|                   | VP, ST, RHMAX, WS | MARS | 0.638 | 0.990 | 0.991    |
|                   | VP, ST, RHMAX, WS | RBFNN | 0.611 | 0.991 | 0.991    |
|                   | VP, ST, RHMAX, TMAX | MLPNN | 0.597 | 0.991 | 0.992    |
|                   | VP, ST, RHMAX, SR | KELM | 0.551 | 0.993 | 0.993    |
Table 4. Cont.

| Input Combination       | Model | Testing Phase |   |   |
|-------------------------|-------|---------------|---|---|
|                         |       | RMSE ($^\circ$C) | NSE | $R^2$ |
| Five-input              |       |               |    |    |
| VP, ST, RHMAX, TMIN, RHMIN | BRT  | 0.604 | 0.991 | 0.991 |
| VP, ST, RHMAX, TMIN, WS  | MARS | 0.643 | 0.990 | 0.990 |
| VP, ST, RHMAX, TMIN, WS  | RBFNN | 0.693 | 0.989 | 0.989 |
| VP, ST, RHMAX, TMIN, WS  | MLPNN | 0.590 | 0.992 | 0.992 |
| VP, ST, RHMAX, TMIN, RHMIN | KELM | 0.535 | 0.993 | 0.993 |
| Six-input               |       |               |    |    |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX | BRT  | 0.603 | 0.991 | 0.991 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX | MARS | 0.638 | 0.990 | 0.990 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX | RBFNN | 0.576 | 0.992 | 0.992 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX | MLPNN | 0.542 | 0.993 | 0.993 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX | KELM | 0.496 | 0.994 | 0.994 |
| Seven-input             |       |               |    |    |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, SR | BRT  | 0.601 | 0.991 | 0.991 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, WS | MARS | 0.655 | 0.990 | 0.990 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, WS | RBFNN | 0.576 | 0.992 | 0.992 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, WS | MLPNN | 0.492 | 0.994 | 0.994 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, WS | KELM | 0.485 | 0.994 | 0.994 |
| Eight-input             |       |               |    |    |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, SR, WS | BRT  | 0.643 | 0.990 | 0.990 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, SR, WS | MARS | 0.634 | 0.990 | 0.990 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, SR, WS | RBFNN | 0.922 | 0.980 | 0.980 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, SR, WS | MLPNN | 0.583 | 0.992 | 0.992 |
| VP, ST, RHMAX, TMIN, RHMIN, TMAX, SR, WS | KELM | 0.492 | 0.994 | 0.994 |

The best models yielding the lowest RMSE are shown in boldface.

By analyzing the four-input composite models, it is clear that the composite of VP, ST, RHMAX, and WS parameters was better than the other four-input composites in the MARS and RBFNN models. Additionally, the BRT model ($RMSE = 0.598 \, ^\circ C$, $NSE = 0.991$, and $R^2 = 0.991$) provided an excellent performance for the composite of VP, ST, RHMAX, and RHMIN parameters, and the MLPNN model ($RMSE = 0.597 \, ^\circ C$, $NSE = 0.991$, and $R^2 = 0.992$) was accurate for the composite of VP, ST, RHMAX, and TMAX parameters. In addition, the KELM model ($RMSE = 0.551 \, ^\circ C$, $NSE = 0.993$, and $R^2 = 0.993$) exhibited the best accuracy with the composite of VP, ST, RHMAX, and SR parameters among the other four-input composite models during the testing phase. Five-input composite models showed that the composite of VP, ST, RHMAX, TMIN, and WS parameters was better than the other five-input composites in the MARS, RBFNN, and MLPNN models. Among three models (i.e., MARS, RBFNN, and MLPNN), the MLPNN model ($RMSE = 0.590 \, ^\circ C$, $NSE = 0.992$, and $R^2 = 0.992$) surpassed the RBFNN and MARS models during the testing phase. Additionally, the BRT ($RMSE = 0.604 \, ^\circ C$, $NSE = 0.991$, and $R^2 = 0.991$) and KELM ($RMSE = 0.535 \, ^\circ C$, $NSE = 0.993$, and $R^2 = 0.993$) models were accurate for the composite of VP, ST, RHMAX, TMIN, and RHMIN parameters. Moreover, the KELM model provided the best performance, based on the five-input composite models during the testing phase. The performance of six-input composite models showed that the composite of VP, ST, RHMAX, TMIN, RHMIN, and TMAX parameters was better than the other six-input composites in the BRT, MARS, and KELM models. Among them, the KELM model ($RMSE = 0.496 \, ^\circ C$, $NSE = 0.994$, and $R^2 = 0.994$) clearly outstripped the BRT and MARS models during the testing phase. In addition, the RBFNN ($RMSE = 0.576 \, ^\circ C$, $NSE = 0.992$, and $R^2 = 0.992$) and MLPNN ($RMSE = 0.542 \, ^\circ C$, $NSE = 0.993$, and $R^2 = 0.993$) models were accurate for the composite of VP, ST, RHMAX, TMIN, RHMIN, and WS parameters. In addition, the KELM model gave the best accuracy based on the six-input composite models during the testing phase. In the case of seven- and eight-input composite models, the predictive results of three evaluation indicators revealed that the composite of VP, ST, RHMAX, TMIN, RHMIN,
TMAX, and WS parameters was better than other seven-input composites for the MARS, RBFNN, MLPNN, and KELM models. Only the BRT model ($RMSE = 0.601 \, ^\circ C$, $NSE = 0.991$, and $R^2 = 0.991$) was accurate for the composite of VP, ST, RHMAX, TMIN, RHMIN, TMAX, and SR parameters compared to the other seven-input composites. The KELM model ($RMSE = 0.485 \, ^\circ C$, $NSE = 0.994$, and $R^2 = 0.994$) obviously obtained the best performance compared to other seven-input composite models during the testing phase. The performance of eight-input composite models showed that the KELM model ($RMSE = 0.492 \, ^\circ C$, $NSE = 0.994$, and $R^2 = 0.994$) furnished better results than the BRT, MARS, RBFNN, and MLPNN models during the testing phase. Considering the best composite models, the best performance of the developed models (i.e., BRT (one-input), MARS (eight-input), RBFNN (three-input), MLPNN, and KELM (seven-input)) can be identified based on different composites of input parameters during the testing phase. It can be seen from Table 4 that the optimized structures of all input composites for the KELM model gave a better performance than those of BRT, MARS, RBFNN, and MLPNN models during the testing phase. Thus, the KELM model was more accurate than the BRT, MARS, RBFNN, and MLPNN models to predict and generate the time series of dew point temperature at UC Riverside station.

Scatter plots of observed and estimated daily $T_{dew}$ are shown in Figure 6 using the best composite models during the testing phase at UC Riverside station. It can be seen from $R^2$ values that there is a trivial difference among the BRT, MARS, RBFNN, MLPNN, and KELM models. In addition, the KELM model gave a better performance than the BRT, MARS, RBFNN, and MLPNN models, while the MARS model provided the lowest accuracy among the best composite models at UC Riverside station.

![Figure 6](image_url). Comparison of observed and estimated daily $T_{dew}$ values using the best composite models (UC Riverside station).
Figure 7 compares the RMSE values for the best composite models during the testing phase at UC Riverside station. It can be seen from Figure 7 that the RMSE values of the BRT, RBFNN, and MARS models were larger than those of the MLPNN and KELM models during the testing phase. In addition, the KELM model accomplished the best accuracy, whereas the BRT model provided the least accuracy, based on the best composite models at UC Riverside station.

Figure 7. Comparison of RMSE values of the best composite models during the testing phase (UC Riverside station).

Figure 8 explains the error histogram comprising mean ($\mu$) and standard deviation ($\sigma$) for the best composite models at UC Riverside station. The comparison displays that the KELM model provided the lowest standard deviation, whereas the MARS model gave the highest standard deviation based on the best composite models during the testing phase. This trails the pattern of RMSE values for the best composite models during the testing phase at UC Riverside station.

Figure 8. Cont.
Figure 8. Error histogram of the best composite models ((a) BRT, (b) MLPNN, (c) RBFNN, (d) MARS, and (e) KELM; UC Riverside station).

5.3. Discussion

The results showed that the best composite models took on the nonlinear behavior of dew point temperature at both stations. A comparison based on RMSE values among the best composite models at Durham station supported that the KELM model improved by 8.5% (BRT model), 3.2% (RBFNN model), 4.1% (MARS model), and 1.9% (MLPNN model) during the testing phase. Additionally, a comparison based on RMSE values among the best composite models at UC Riverside station showed that the KELM model enhanced the accuracy by 18.1% (BRT model), 23.5% (MARS model), 15.2% (RBFNN model), and 1.4% (MLPNN model) during the testing phase.

Within the category of the best composite models, the predictive accuracy using the KELM model at UC Riverside station improved over that of the other models, whereas the predictive accuracy at Durham station was slightly enhanced. The improvement difference between the two stations came from the characteristics of the data available. This can be found from previous research [19,45,46]. If two (or three) potential models provided the best predictive accuracy, additional approaches (e.g., null hypothesis [47] and Akaike’s information criterion [48]) were recommended in order to select the best model. The authors of [8] developed the GRNN and MLP models for predicting daily dew point temperature at Durham and UC Riverside stations. They showed that the best model (i.e., the GRNN4 model) provided RMSE = 0.07 °C (Durham station) and 0.08 °C (UC Riverside station). Reference [20] employed the GRNN, SOM, ANFIS-C, and ANFIS-G models for predicting daily dew point temperature at Daegu, Pohang, and Ulsan stations, South Korea. The authors found that GRNN, ANFIS-C, and ANFIS-G models were more accurate than the SOM model. In addition, the predictive results followed the previous studies of [8,20].

In addition, different nature-inspired optimization algorithms and data preprocessing tools can be combined with machine learning models to enhance the predictive accuracy of the applied models. To reinforce the model accuracy of this study, continuous research utilizing the different machine learning models, evolutionary algorithms, and data preprocessing techniques should be undertaken for predicting dew point temperature.

6. Conclusions

This study was intended to investigate a kernel extreme learning machine (KELM) model to estimate daily dew point temperature (T_{dew}) at two different stations in the USA. The KELM model was trained to estimate daily T_{dew} by employing hydroclimatic variables, including WS, TMAX, TMIN, RHMAX, RHMINT, VP, ST, and SR, as inputs. Additionally, eight different scenarios were applied to investigate the effect of hydrometeorological variables on daily T_{dew} estimation using different machine learning models. The KELM models were compared with the BRT, MARS, RBFNN, and MLPNN models with respect to the Nash–Sutcliffe efficiency (NSE) statistical indices and coefficient of determination ($R^2$) and root mean square error (RMSE) indicators. The KELM models using the
three-input parameters of VP, TMAX, and RHMIN and the seven-input parameters of VP, ST, RHMAX, TMIN, RHMIN, TMAX, and WS outperformed the other models for the estimation of daily \( T_{dew} \) at Durham and UC Riverside stations, respectively. The results confirmed that the KELM model improved by 8.5% (BRT model), 3.2% (RBFNN model), 4.1% (MARS model), and 1.9% (MLPNN model) based on the RMSE values among the best composite models during the testing phase at Durham station. In addition, the KELM model enhanced the accuracy by 18.1% (BRT model), 23.5% (MARS model), 15.2% (RBFNN model), and 1.4% (MLPNN model) based on the RMSE values among the best composite models during the testing phase in UC Riverside station. The results suggest that the KELM model can be successfully used for estimating dew point temperature using weather data as an important parameter for sustainable water resource management.

**Author Contributions:** Conceptualization and data analysis, M.A.; methodology, M.Z.-K.; software, M.A.; funding acquisition, N.W.K.; writing—original draft preparation, S.H., M.Z.-K., N.W.K., and S.K.; visualization, M.A.; supervision, M.A.; writing—review and editing, S.K., M.A., and V.P.S.; project administration, M.A. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Korea Institute of Civil Engineering and Building Technology, grant number 20200027-001.

**Acknowledgments:** The authors wish to express our gratitude to the two anonymous reviewers whose suggestions and remarks have greatly helped us to improve the quality of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**

1. Mahmood, R.; Hubbard, K.G. Assessing bias in evapotranspiration and soil moisture estimates due to the use of modeled solar radiation and dew point temperature data. *Agric. For. Meteorol.* **2005**, *130*, 71–84. [CrossRef]
2. Zounemat-Kermani, M. Hourly predictive Levenberg–Marquardt ANN and multi linear regression models for predicting of dew point temperature. *Meteorol. Atmos. Phys.* **2012**, *117*, 181–192. [CrossRef]
3. Garcia, M.; Raes, D.; Jacobsen, S.E. Evapotranspiration analysis and irrigation requirements of quinoa (*Chenopodium quinoa*) in the Bolivian highlands. *Agric. Water Manag.* **2014**, *7*, 5037–5050. [CrossRef]
4. Mortuza, M.R.; Selmi, S.; Khudri, M.M.; Ankur, A.K.; Rahman, M.M. Evaluation of temporal and spatial trends in relative humidity and dew point temperature in Bangladesh. *Arab. J. Geosci.* **2014**, *7*, 5037–5050. [CrossRef]
5. Ali, H.; Fowler, H.J.; Mishra, V. Global observational evidence of strong linkage between dew point temperature and precipitation extremes. *Geophys. Res. Lett.* **2018**, *45*, 12–320. [CrossRef]
6. Bui, A.; Johnson, F.; Wasko, C. The relationship of atmospheric air temperature and dew point temperature to extreme rainfall. *Environ. Res. Lett.* **2019**, *14*, 074025. [CrossRef]
7. Shank, D.B.; Hoogenboom, G.; McClendon, R.W. Dew point temperature prediction using artificial neural networks. *J. Appl. Meteorol. Climatol.* **2008**, *47*, 1757–1769. [CrossRef]
8. Kim, S.; Singh, V.P.; Lee, C.J.; Seo, Y. Modeling the physical dynamics of daily dew point temperature using soft computing techniques. *KSCE J. Civ. Eng.* **2015**, *19*, 1930–1940. [CrossRef]
9. Nadig, K.; Potter, W.; Hoogenboom, G.; McClendon, R. Comparison of individual and combined ANN models for prediction of air and dew point temperature. *Appl. Intell.* **2013**, *39*, 354–366. [CrossRef]
10. Mohammadi, K.; Shamshirband, S.; Motamed, S.; Petković, D.; Hashim, R.; Gocic, M. Extreme learning machine based prediction of daily dew point temperature. *Comput. Electron. Agric.* **2015**, *117*, 214–225. [CrossRef]
11. Alizamir, M.; Kim, S.; Kisi, O.; Zounemat-Kermani, M. Deep echo state network: A novel machine learning approach to model dew point temperature using meteorological variables. *Hydrol. Sci. J.* **2020**, *10*, 18. [CrossRef]
12. Dong, J.; Wu, L.; Liu, X.; Li, Z.; Gao, Y.; Zhang, Y.; Yang, Q. Estimation of daily dew point temperature by using bat algorithm optimization based extreme learning machine. *Appl. Therm. Eng.* **2020**, *165*, 114569. [CrossRef]
13. Shiri, J. Prediction vs. estimation of dew point temperature: Assessing GEP, MARS and RF models. *Hydrol. Res.* **2019**, *50*, 633–643. [CrossRef]
14. Qasem, S.N.; Samadianfard, S.; Nahand, H.S.; Mosavi, A.; Shamshirband, S.; Chau, K.W. Estimating daily dew point temperature using machine learning algorithms. *Water* 2019, 11, 582. [CrossRef]

15. Naganna, S.R.; Deka, P.C.; Ghorbani, M.A.; Biazar, S.M.; Al-Ansari, N.; Yaseen, Z.M. Dew point temperature estimation: Application of artificial intelligence model integrated with nature-inspired optimization algorithms. *Water* 2019, 11, 742. [CrossRef]

16. Attar, N.F.; Khalili, K.; Behmanesh, J.; Khammohammadi, N. On the reliability of soft computing methods in the estimation of dew point temperature: The case of arid regions of Iran. *Comput. Electron. Agric.* 2018, 153, 334–346. [CrossRef]

17. Mehdizadeh, S.; Behmanesh, J.; Khalili, K. Application of gene expression programming to predict daily dew point temperature. *Appl. Therm. Eng.* 2017, 112, 1097–1107. [CrossRef]

18. Deka, P.C.; Patil, A.P.; Yeswanth Kumar, P.; Naganna, S.R. Estimation of dew point temperature using SVM and ELM for humid and semi-arid regions of India. *ISH J. Hydraul. Eng.* 2018, 24, 190–197. [CrossRef]

19. Shiri, J.; Kim, S.; Kisi, O. Estimation of daily dew point temperature using genetic programming and neural networks approaches. *Hydrolog. Res.* 2014, 45, 165–181. [CrossRef]

20. Kisi, O.; Kim, S.; Shiri, J. Estimation of dew point temperature using neuro-fuzzy and neural network techniques. *Theor. Appl. Climatol.* 2013, 114, 365–373. [CrossRef]

21. Al-Shammari, E.T.; Mohammadi, K.; Keivani, A.; Ab Hamid, S.H.; Akib, S.; Shamshirband, S.; Petković, D. Prediction of daily dew point temperature using a model combining the support vector machine with firefly algorithm. *J. Irrig. Drain. Eng.* 2016, 142, 04016013. [CrossRef]

22. Amirmojahedi, M.; Mohammadi, K.; Shamshirband, S.; Danesh, A.S.; Mostafaiepour, A.; Kamsin, A. A hybrid computational intelligence method for predicting dew point temperature. *Environ. Earth Sci.* 2016, 75, 415. [CrossRef]

23. Baghban, A.; Bahadori, M.; Rozyn, J.; Lee, M.; Abbas, A.; Bahadori, A.; Rahimali, A. Estimation of air dew point temperature using computational intelligence schemes. *Appl. Therm. Eng.* 2016, 93, 1043–1052. [CrossRef]

24. Zounemat-Kermani, M.; Kisi, O.; Rajaei, T. Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff. *Appl. Soft Comput.* 2013, 13, 4633–4644. [CrossRef]

25. Ghanbari, A.; Kardani, M.N.; Moazami Goodarzi, A.; Janghorban Lariche, M.; Baghban, A. Neural computing approach for estimation of natural gas dew point temperature in glycol dehydration plant. *Int. J. Ambient Energy* 2018, 1–8. [CrossRef]

26. Zounemat-Kermani, M.; Kisi, O.; Rajaei, T. Performance of radial basis and LM-feed forward artificial neural networks for predicting daily watershed runoff. *Appl. Soft Comput.* 2013, 13, 4633–4644. [CrossRef]

27. Nacar, S.; Hınıs, M.A.; Kankal, M. Forecasting daily streamflow discharges using various neural network models and training algorithms. *KSCJ J. Civ. Eng.* 2018, 22, 3676–3685. [CrossRef]

28. Kisi, O.; Alizamir, M.; Gorgij AR, D. Dissolved oxygen prediction using a new ensemble method. *Environ. Sci. Pollut. Res.* 2020, 1–15. [CrossRef]

29. Alizamir, M.; Kisi, O.; Muhammad Adnan, R.; Kuriqi, A. Modelling reference evapotranspiration by combining neuro-fuzzy and evolutionary strategies. *Acta Geophys.* 2020, 117239. [CrossRef]

30. Alizamir, M.; Kisi, O.; Ahmed, A.N.; Mert, C.; Fai, C.M.; Kim, S.; Kim, N.W.; El-Shafie, A. Advanced machine learning model for better prediction accuracy of soil temperature at different depths. *PLoS ONE* 2020, 15, e0231055. [CrossRef]

31. Alizamir, M.; Kisi, O.; Muhammad Adnan, R.; Kuriqi, A. Modelling reference evapotranspiration by combining neuro-fuzzy and evolutionary strategies. *Acta Geophys.* 2020, 1–14. [CrossRef]

32. Zhang, G.P.; Qi, M. Neural network forecasting for seasonal and trend time series. *Eur. J. Oper. Res.* 2005, 160, 501–514. [CrossRef]

33. Parsaie, A.; Najafian, S.; Shamsi, Z. Predictive modeling of discharge of flow in compound open channel using radial basis neural network. *Modeling Earth Syst. Environment.* 2016, 2, 150. [CrossRef]

34. Silva, D.; Nunes, I.; Spatti, D.H.; Flauzino, R.A.; Liboni, L.H.B.; Alves, S.F.d.R. Artificial Neural Networks; Springer International Publishing: Cham, Switzerland, 2017; p. 39.

35. Huang, G.-B.; Zhu, Q.-Y.; Siew, C.-K. Extreme learning machine: A new learning scheme of feedforward neural networks. *IEEE Int. J. Conf. Neural Netw.* 2004, 2, 985–990.
36. Alizamir, M.; Kisi, O.; Zounemat-Kermani, M. Modelling long-term groundwater fluctuations by extreme learning machine using hydro-climatic data. *Hydrol. Sci. J.* 2018, 63, 63–73. [CrossRef]
37. Adnan, R.M.; Liang, Z.; Trajkovic, S.; Zounemat-Kermani, M.; Li, B.; Kisi, O. Daily streamflow prediction using optimally pruned extreme learning machine. *J. Hydrol.* 2019, 577, 123981. [CrossRef] [PubMed]
38. Huang, G.B.; Zhou, H.M.; Ding, X.J.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. Part B-Cybern.* 2012, 42, 513–529. [CrossRef] [PubMed]
39. Kisi, O.; Alizamir, M. Modelling reference evapotranspiration using a new wavelet conjunction heuristic method: Wavelet extreme learning machine vs wavelet neural networks. *Agric. For. Meteorol.* 2018, 263, 41–48. [CrossRef]
40. Deo, R.C.; Kisi, O.; Singh, V.P. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmos. Res.* 2017, 184, 149–175. [CrossRef] [PubMed]
41. Emamgolizadeh, S.; Bateni, S.M.; Shahsavani, D.; Ashrafi, T.; Ghorbani, H. Estimation of soil cation exchange capacity using genetic expression programming (GEP) and multivariate adaptive regression splines (MARS). *J. Hydrol.* 2015, 529, 1590–1600. [CrossRef]
42. Kuter, S.; Akyurek, Z.; Weber, G.W. Retrieval of fractional snow covered area from MODIS data by multivariate adaptive regression splines. *Remote Sens. Environ.* 2018, 205, 236–252. [CrossRef]
43. Elith, J.; Leathwick, J.R.; Hastie, T. A working guide to boosted regression trees. *J. Anim. Ecol.* 2008, 77, 802–813. [CrossRef] [PubMed]
44. Persson, C.; Bacher, P.; Shiga, T.; Madsen, H. Multi-site solar power forecasting using gradient boosted regression trees. *Sol. Energy* 2017, 150, 423–436. [CrossRef]
45. Tokar, A.S.; Johnson, P.A. Rainfall-runoff modeling using artificial neural networks. *J. Hydrol. Eng.* 1999, 4, 232–239. [CrossRef]
46. Kim, S.; Kim, H.S. Uncertainty reduction of the flood stage forecasting using neural networks model. *J. Am. Water Resour. Assoc.* 2008, 44, 148–165. [CrossRef]
47. McCuen, R.H. Microcomputer Applications in Statistical Hydrology; Eaglewood Cliffs, N.J., Ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 1993.
48. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 1974, 19, 716–723. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).