

Article

Application of Advanced Optimized Soft Computing Models for Atmospheric Variable Forecasting

Rana Muhammad Adnan ¹, Sarita Gajbhiye Meshram ², Reham R. Mostafa ³,
Abu Reza Md. Towfiqul Islam ^{4,*}, S. I. Abba ⁵, Francis Andorful ⁶ and Zhihuan Chen ^{7,*}

¹ School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China

² Water Resources and Applied Mathematics Research Lab, Nagpur 440027, India

³ Information Systems Department, Faculty of Computers and Information Sciences, Mansoura University, Mansoura 35516, Egypt

⁴ Department of Disaster Management, Begum Rokeya University, Rangpur 5400, Bangladesh

⁵ Interdisciplinary Research Center for Membranes and Water Security, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia

⁶ Department of Geography and Resource Development, University of Ghana, Accra 23321, Ghana

⁷ Engineering Research Center for Metallurgical Automation and Measurement Technology of Ministry of Education, Wuhan University of Science and Technology, Wuhan 431400, China

* Correspondence: towfiq_dm@brur.ac.bd (A.R.M.T.I.); czh@wust.edu.cn (Z.C.)

Abstract: Precise Air temperature modeling is crucial for a sustainable environment. In this study, a novel binary optimized machine learning model, the random vector functional link (RVFL) with the integration of Moth Flame Optimization Algorithm (MFO) and Water Cycle Optimization Algorithm (WCA) is examined to estimate the monthly and daily temperature time series of Rajshahi Climatic station in Bangladesh. Various combinations of temperature and precipitation were used to predict the temperature time series. The prediction ability of the novel binary optimized machine learning model (RVFL-WCAMFO) is compared with the single optimized machine learning models (RVFL-WCA and RVFL-MFO) and the standalone machine learning model (RVFL). Root mean square errors (RMSE), the mean absolute error (MAE), the Nash–Sutcliffe efficiency (NSE), and the determination coefficient (R^2) statistical indexes were utilized to access the prediction ability of the selected models. The proposed binary optimized machine learning model (RVFL-WCAMFO) outperformed the other single optimized and standalone machine learning models in prediction of air temperature time series on both scales, i.e., daily and monthly scale. Cross-validation technique was applied to determine the best testing dataset and it was found that the M3 dataset provided more accurate results for the monthly scale, whereas the M1 dataset outperformed the other two datasets on the daily scale. On the monthly scale, periodicity input was also added to see the effect on prediction accuracy. It was found that periodicity input improved the prediction accuracy of the models. It was also found that precipitation-based inputs did not provided very accurate results in comparison to temperature-based inputs. The outcomes of the study recommend the use of RVFL-WCAMFO in air temperature modeling.

Keywords: machine learning; hybrid modeling; moth flame optimization (MFO); water cycle algorithm (WCA); random vector functional link (RVFL)

MSC: 68U01



Citation: Adnan, R.M.; Meshram, S.G.; Mostafa, R.R.; Islam, A.R.M.T.; Abba, S.I.; Andorful, F.; Chen, Z. Application of Advanced Optimized Soft Computing Models for Atmospheric Variable Forecasting. *Mathematics* **2023**, *11*, 1213. <https://doi.org/10.3390/math11051213>

Academic Editors: Adrian Deaconu, Petru Adrian Cotfas, Daniel Tudor Cotfas and Andrea Scozzari

Received: 19 December 2022

Revised: 17 February 2023

Accepted: 27 February 2023

Published: 1 March 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The global surface temperature (GT) has grown since the preindustrial era in the years 2001 to 2020 by 0.99 °C in comparison to the years 1850 to 1900 [1]. In recent years, there has been a clear trend of increasing frequency of extreme heat events around the world, with many regions experiencing more frequent and severe heatwaves. This trend is particularly evident in eastern Africa, India, and the Amazon basin. These heatwaves have

led to record-breaking temperatures, and have had significant impacts on human health, agriculture, and ecosystems.

One example of this trend is the record-breaking heatwave that occurred in France in June 2019, where the highest temperature ever recorded was 46.0 °C. However, just a couple of years later, in August 2021, Sicily, Italy broke that record with a temperature of 48.8 °C. This highlights the increasing severity and frequency of heatwaves in recent years, and the need for action to address the underlying causes of these events and to prepare for the impacts they are likely to have in the future [2].

Climate change is a main factor in these heatwaves, which are caused by the warming of the planet due to the increasing concentrations of greenhouse gases in the atmosphere. These heatwaves are likely to become more frequent and severe in the future unless we take action to reduce our greenhouse gas emissions and implement other mitigation and adaptation strategies. As our planet faces the challenges of climate change, it is crucial that we have access to accurate weather data for urban areas. This information is essential for both the administration and planning of cities, as well as for the well-being of the people who live in them. Having accurate weather data allows city officials to make informed decisions about infrastructure and resource allocation, such as identifying areas at risk of flooding and implementing measures to reduce the impacts of extreme weather events. It also allows them to plan for the future, by taking into account the potential impacts of climate change on the city and its residents. For city residents, accurate weather data can help them to stay safe and prepare for extreme weather events. It can also be used to make decisions about daily activities, such as whether to walk or bike to work or if it is necessary to use public transportation [3].

Air temperature is a vital meteorological element that plays a crucial role in shaping the growth and productivity of crops. It is one of the most important factors that determine the suitability of a region for agriculture, and its fluctuations can have a significant impact on crop development and yield. Temperature also affects other environmental elements such as air pressure, relative humidity, wind speed, and rainfall. These factors interact with each other and can have a cascading effect on the growth and productivity of crops. For example, high temperatures can increase evaporation rates and reduce the availability of water for plants, while low temperatures can slow down plant growth and reduce crop yields. In addition to its direct impact on crop growth, air temperature also plays a role in the distribution and survival of pests and diseases. High temperatures can increase the growth rate of some pests and diseases, while low temperatures can slow them down or even kill them [4]. Extreme temperatures can have a variety of detrimental effects on individuals, communities and the environment, leading to a wide range of thermal disasters such as, health effects (such as heatstroke), catastrophic crop failures, wildfires, and power outages [5].

According to recent studies, air temperature may be related to the development of thrombus, which is a blood clot that can happen in the veins. The association between air temperature and venous thromboembolism (VTE), a disorder that develops when a blood clot forms in a vein and can cause major health issues such as deep vein thrombosis and pulmonary embolism, has specifically been the subject of multiple research studies [6]. According to these studies, there may be a link between high air temperatures and a higher incidence of VTE. This is thought to be caused by how temperature affects the biological processes that regulate blood clot formation [7].

Africa's socioeconomic development, agriculture, and water security are also affected by shifting rainfall patterns and rising air temperatures [8]. Climate change's altered rainfall patterns can result in droughts and floods that could have a serious impact on agriculture and food security [9]. Crop production can be decreased by droughts, and crops and infrastructure can be destroyed by floods. Increased evaporation rates and decreased soil moisture due to rising air temperatures, which are also a result of climate change, can put additional strain on the agricultural industry and cause crop failures, yield losses, and the spread of pests and diseases that can destroy crops. A considerable impact on human

populations as well as water security may result from these shifting rainfall patterns and rising air temperatures.

There are numerous ways to measure the temperature of the air over time and in different locations [10]. To accurately monitor air temperature, a ground-based thermometer (2 m above ground) with sufficient accuracy and temporal resolution is typically employed. These datasets, however, do not adequately represent the vast variety of surfaces because they were collected as point samples [11]. To improve the accuracy of air temperature estimation, various techniques can be employed, such as using output-improving machine learning algorithms. These algorithms can take into account a variety of data sources, such as satellite images, remote sensing data, or weather station data. By using these techniques, it is possible to improve the accuracy of temperature estimations and to create more detailed and reliable temperature maps [12].

Additionally, using a combination of ground-based and satellite-based measurements can provide a more comprehensive understanding of temperature variations across different surfaces, and can help to identify patterns and trends that are not visible when using point-based measurements alone [13–17]. However, acquiring more input data to precisely estimate the temperature is a challenging task. Therefore, researchers prefer to use less complex models with fewer inputs to predict temperature.

Many different ML-based algorithms have been studied for use in forecasting temperature with minimum input data. The most often used techniques in the study of air temperature time series are Artificial Neural Networks (ANN) and Support Vector Machines (SVM). In particular, Multi-Layer Perceptron Neural Networks (MLPNN) and Radial Basis Function Neural Networks (RBFNN) are the most common ANN models used to predict temperature values [18–24]. The most popular optimization algorithms are Radial Function Base Kernels which are used in the majority of SVM model-related publications [25–28]. Robert et al. [29] predict the short time air temperature using the SVM model. They found that SVM provided more accurate results in comparison of the ANN model for one hour ahead temperature predictions. Salcedo-Sanz et al. [30] forecasted the air temperature of Australia and New Zealand on a long time scale. They predicted the mean monthly air temperature using SVM and MLPNN models and found that SVM outperformed the MLPNN models.

In the above discussed literature, researchers predict the maximum, minimum, and average temperature on a short scale (hourly and daily) and a long scale (weekly and monthly). However, optimal selection of control parameters in machine learning model is a challenging task and can aid in improving the predicted results. Hybrid machine learning algorithms are a combination of two or more machine learning algorithms that are used together to create a more powerful and accurate model. These algorithms are used to solve complex problems that require a huge amount of input data. Venkadesh et al. [31] utilized the hybrid machine learning models to predict the air temperatures on a short time scale. They used the ANN hybrid model based on genetic algorithm (GA) to forecast one hour ahead air temperature. They found that the hybrid ANN model produced more precise results than standalone models. Azad et al. [32] applied the adaptive neuro-fuzzy inference system (ANFIS) based hybrid models to forecast the air temperature on a long scale. They utilized genetic algorithm (GA), particle swarm optimization (PSO), ant colony optimization for continuous domains (ACO_R), and differential evolution (DE) to optimize the control parameters of the ANFIS model. They found that hybrid models outperformed the standalone ANFIS models.

It was discovered in the aforementioned discussion that hybrid machine learning models outperformed independent machine learning models in terms of results. However, because of their non-stationary, stochastic character with data noise and modeling of hydrological variables such as air temperature, there is still room to improve the accuracy and time of computation. Temperature forecasting is a challenging task due to other atmospheric parameters' (wind speed, humidity, air vapor pressure) effects on this. However, acquiring such a huge amount of data is a difficult and costly task. In hybrid models,

the optimization algorithm faces strong exploration or exploitation abilities in searching control parameters. Hybrid models are more effective because they can capitalize on each algorithm's advantages and combine them to produce a more potent model. Several models need to be combined to improve the results of non-stationary and stochastic data such as air temperature. The model presented here ensures excellent accuracy in calculating the air temperature while requiring less computing time. Therefore, more durable machine learning models are required to capture nonlinear trends in hydrological variables. According to research, a good optimization algorithm should be able to balance its exploitation and exploration capabilities. Due to the nonlinear stochastic character of the variable and the difficulties in predicting it as a result of a variety of external inputs, it is difficult to achieve this equilibrium for a single optimization in the case of modeling hydrological variable time series. By utilizing the stronger (exploring/exploitation) aspects of another optimization algorithm, this study closes this gap by bolstering the weaker (exploring/exploitation) aspects of one optimization technique. This prompts us to combine the Water Cycle Method (WCA) with the Moth Flame Optimization (MFO) algorithm to improve its capacity for exploitation.

In particular, a new model is presented that ensures excellent accuracy in calculating air temperature while requiring less computing time. The model uses an optimization algorithm that balances its skills for exploration and exploitation, which is difficult to achieve for a single optimization in the case of hydrological variable time series modeling due to the nonlinear and stochastic nature of the variable. To improve the capacity for exploitation, the Water Cycle Method (WCA) is combined with the Moth Flame Optimization (MFO) algorithm.

In this study, the authors present a new method for improving the prediction of air temperature by combining a hybrid random vector functional link network (RVFL) with a heuristic optimization technique called Water Cycle-Moth Flame Optimization (WCAMFO). The method is compared to other RVFL-based approaches, such as RVFL-WCA, RVFL-MFO, RVFL-WCAMFO, and single RVFL, in order to evaluate its performance. In Section 2, the authors discuss the most popular machine-learning-based techniques and their related topics. They also present the RVFL model and its implementation with the WCAMFO algorithm, which is designed to improve the accuracy of temperature predictions. In Section 3, the authors present the results and discussion of the comparison of the different methods. They evaluate the performance of the proposed method in terms of prediction accuracy and computational efficiency. Finally, in Section 4, the authors discuss the findings of the study and identify research gaps in temperature forecasting. They conclude that the proposed method, RVFL paired with WCAMFO, is a promising approach for improving the prediction air temperature. They also suggest that further research is needed to further improve the performance of the model and to make it more widely applicable.

2. Case Study

The climate of the study area is subhumid, warm, and subtropical (Figure 1). The western part of Bangladesh encompasses approximately 41%, or 60,165 km² of the country [33]. The subtropical monsoon climate is experienced by three distinct seasons: winter (Nov–Feb), which is characterized by being cool and dry with almost no rainfall; the pre-monsoon (Mar–May), characterized by being hot and dry; and the monsoon (Jun–Oct), characterized by heavy rainfall. The annual rainfall over the last 30 years was 1600 mm, less than the national average of 2550 mm in the study area [34]. The annual rainfall and average temperature in the study area range from 1492 to 2766 mm, with an average of 1925 mm, and 24.18 to 26.17 °C, with an average of 25.44 °C, respectively [35]. Two distinct landforms characterize the study region. One is the Barind Tract, dissected and undulating, and the other is the floodplains. Geologically, the area consists of the stream and inter-stream recent and Pleistocene sediments that overlie the Gondwana sediments. In this study, daily and monthly average data of temperature variable for the Rajshahi climatic station is collected from the Bangladesh Meteorological Department (BMD) for

the duration of 27 years (1992 to 2018). For data modeling, three equally spaced testing datasets of M1, M2, and M3 using nine years of data.

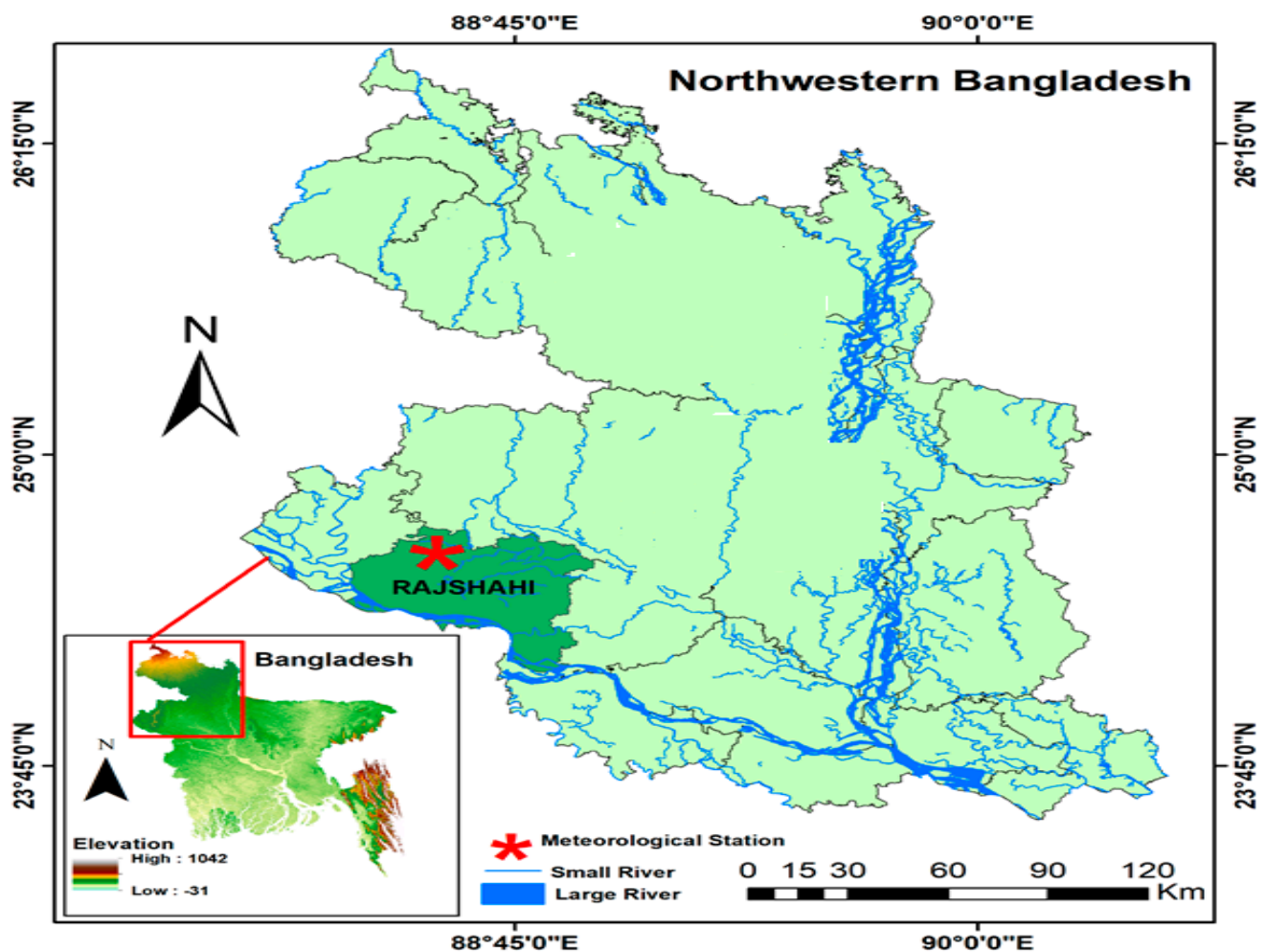


Figure 1. Study area Location.

3. Methods

3.1. Random Vector Functional Link Network (RVFL)

The Random vector functional link (RVFL) network (Figure 2), a modified version of the conventional single-layer feed-forward neural network, was introduced through a successful learning process [36,37]. Input and output layers are connected directly in the RVFL network's general layout, which also includes a layer of nodes known as enhancement neurons that serves as the network's hidden layer. The following is a summary of the RVFL network's fundamental principles:

- i. Overfitting problems are avoided and RVFL's performance is improved because of the layer-by-layer interaction between the input and output layers.
- ii. With randomly selected input weights that are not subject to modification during the training phase, the computational cost is reduced.
- iii. Ridge regression [38] or the Moore-Penrose pseudo-inverse are used to produce the output weights, which are adjusted by traditional learning techniques [37].
- iv. The main advantage of RVFL is that it offers a potential remedy for several critical problems with the conventional learning algorithm in the multi-layer neural network design (poor convergence speed and large computing weight).
- v. In order for the RVFL to function, m training samples of data are fed into the network, with each sample denoted by (x_i, y_i) , where x_i and y_i stand for the input

and output, respectively. After that, the input is transmitted to the enhancement nodes, and the final result is calculated as follows [39]:

$$O_j(a_j x_i + \beta_j) = \frac{1}{1 + e^{-(a_j x_i + \beta_j)}}, \beta_j \in [0, \mathcal{S}], a_j \in [-\mathcal{S}, \mathcal{S}] \quad (1)$$

where a_j represent the weights of the input layer and enhancement nodes. The bias is denoted by β_j , while the scale factor is denoted by $\mathcal{S}$. In conclusion, the output of the RVFL network can be stated as follows:

$$Z = Bw \quad (2)$$

where b is a matrix containing the data from b_1 input data, and b_2 result from the enhancement node

$$B_1 = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{Nn} \end{bmatrix}, B_2 = \begin{bmatrix} \mathcal{O}_1(a_1 x_1 + d_1) & \cdots & \mathcal{O}_p(a_p x_1 + \beta_p) \\ \vdots & \ddots & \vdots \\ \mathcal{O}_1(a_1 x_N + d_1) & \cdots & \mathcal{O}_p(a_p x_N + \beta_p) \end{bmatrix} \quad (3)$$

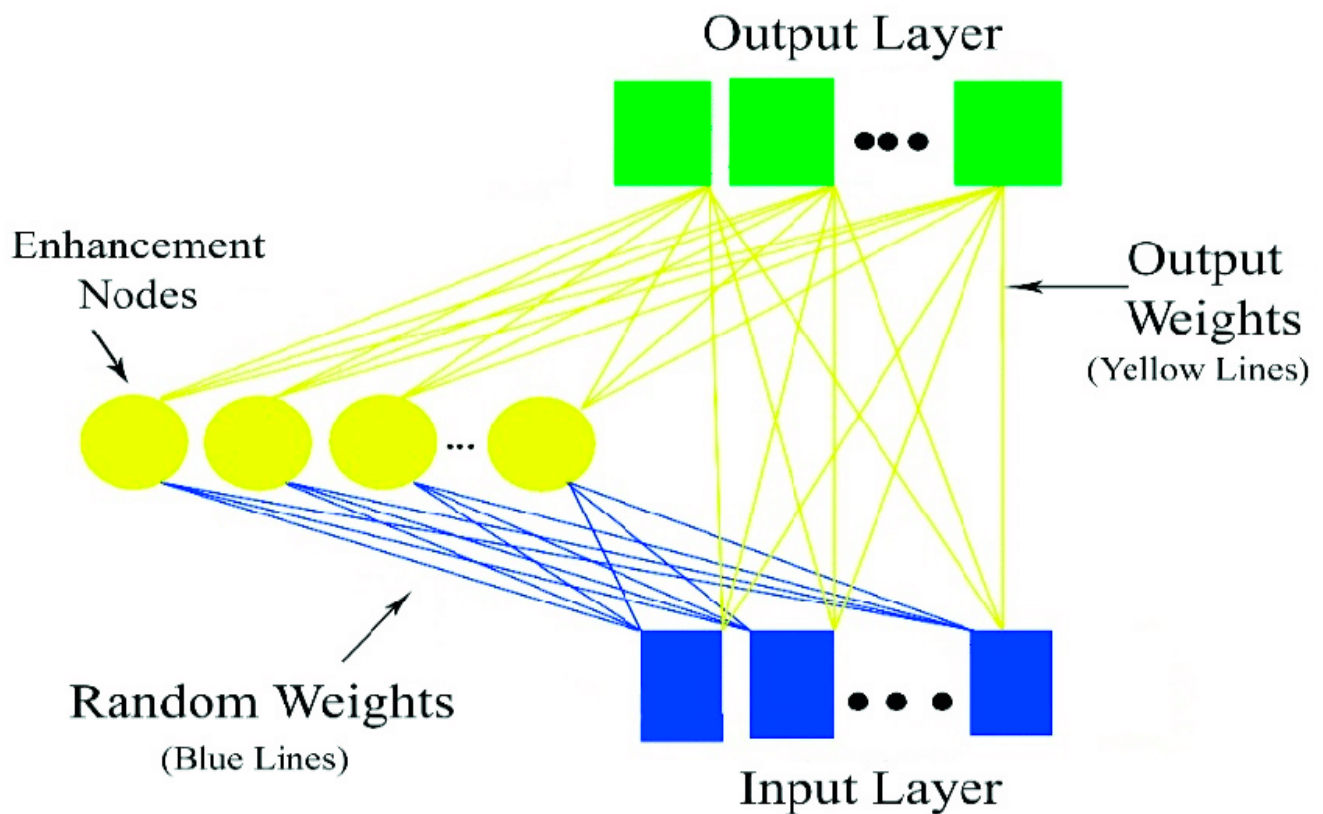


Figure 2. Schematic structure of RVFL.

Through the use of the Moore-Penrose pseudo-inverse, the weight w can be calculated:

$$w = B^\dagger Z \quad (4)$$

where $\dagger$ is the Moore-Penrose pseudo-inverse.

3.2. Moth-Flame Optimization Algorithm (MFO)

The new biologically-inspired optimization method known as the MFO algorithm is based on the flight pattern of a moth, which spirals toward and then clings to an arti-

cial light source to achieve optimization [40]. The essential premise of MFO algorithm (Figure 3) is as follows.

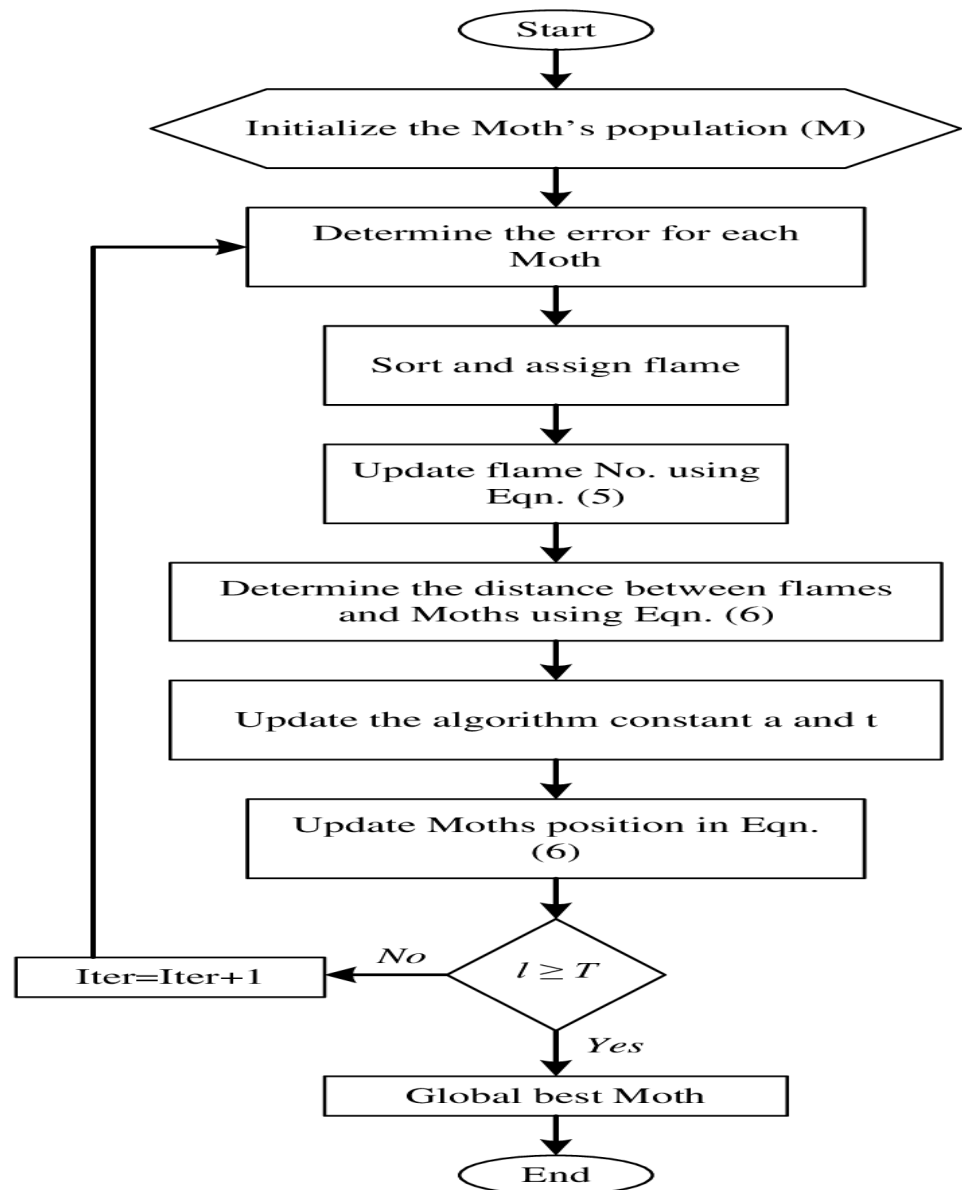


Figure 3. Moth-flame optimization algorithm flowchart.

Assume that the matrices $\mathcal{M} = (\mathcal{N}, dim)$ and $\mathcal{F} = (\mathcal{N}, dim)$ are used to represent the set of moth and flame, and that flames and moths are potential solutions to issues. In the formula, $\mathcal{N}$ represents the maximum number of flames and dim represents the population of moths. The flames reflect the best location obtained from the moths, and the moths represent the real search topics travelling through the scope of the search. According to the greedy retention principle, which stipulates that the flame structure preserves the best $\mathcal{N}$ solutions to the moth's flight history, the position of the moth and the flame structure are updated with the moth's spiral flight [41]. The spiral flying mechanism has the following characteristics:

- (1) The search area must include the spiral flying space.
- (2) The moth serves as the origin of spiral flying.
- (3) The flame position is the spiral's conclusion.

To update each moth's location and replicate its flight mode, use Equation (5).

$$\mathcal{M}_i = d_i * e^{bt} * \cos(2\pi t) + \mathcal{F}_j \quad (5)$$

$\mathcal{M}_i$ and $\mathcal{F}_j$ in the equation stand for the i^{th} moth and the j^{th} flame, respectively. d_i is the distance between them, and its expression is given in Equation (6). The spiral b flight shape is defined by b ($b = 1$) is the spiral flight coefficient, and Equation (7) shows its t expression. Its value range is $[-1, 1]$.

$$d_i = |\mathcal{F}_j - \mathcal{M}_i| \quad (6)$$

$$t = \left(-\frac{\ell + \mathcal{L}}{\mathcal{L}} - 1 \right) \times rand + 1 \quad (7)$$

where $\mathcal{L}$ and ℓ stand for the number of iterations' maximum and current durations. A random number with a uniform $\mathcal{L}$ rand distribution is called rand. When $t = -1$, it indicates that the distance between the flame and the moth is at its shortest, and when $t = 1$, it indicates that it is at its greatest.

Equation (8) is utilized to minimize the number of flames during each iteration in an adaptive linear way, increasing the moth's capacity to mine locally in subsequent iterations. At the start of the iteration, there are N flames. The moths only adjust their position around the best flame at the end of iteration, greatly enhancing the search capabilities of the algorithm.

$$flame_{no} = round\left(\mathcal{N} - \ell \times \frac{\mathcal{N} - 1}{\mathcal{L}}\right) \quad (8)$$

Numerous MFO variants have been developed, including Lévy-flight moth-flame optimization (LMFO) [42], non-dominated sorting moth flame optimization (NS-MFO) [43], enhanced moth-flame optimization (EMFO) [44], water cycle-moth-flame optimization (WCMFO) [45], and sine-cosine moth-flame optimization (SMFO) [46]. MFO and its derivatives still have some flaws, such as poor population diversity [47], premature convergence, local optima trapping [48], and an imbalance between exploration and exploitation [49], which prevent them from meeting the needs of the optimization process for difficult problems. The majority of moths are stuck in the local optima in the early iterations, leading to poor population diversity, which is the main cause of these MFOs' limitations [49].

3.3. Water Cycle Optimization Algorithm (WCA)

Eskandar et al. [50] created the WCA in accordance with the natural water cycle or hydrological cycle. The WCA technique begins with the starting population, referred to as raindrops, like other meta-heuristic algorithms. The initial assumption is that there is rain or other precipitation. The best individual (a water drop) is picked to represent the sea. The remaining raindrops are then regarded as streams that empty into rivers and seas, whereas some of the nice droplets are regarded as rivers. A single solution is referred to as a "raindrop" in the WCA method. Such an array is referred to as a "chromosome" in the GA method. A raindrop is an $\mathcal{X}_{\mathcal{N}_{var}}$ array in a problem of multidimensional optimization, defined as Equation (9).

$$Raindrop = [\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_{\mathcal{N}_{var}}] \quad (9)$$

where $\mathcal{X}_1$ to $\mathcal{X}_{\mathcal{N}_{var}}$ represent the decision variables. To begin with, a sample of the raindrop matrix with the size of $\mathcal{N}_{pop} \times \mathcal{N}_{var}$ is randomly generated.

$$\begin{aligned} Population\ of\ raindrops &= [bmatrix \mathcal{X}_{Raindrop1} Raindrop2; RaindropN pop bmatrix \mathcal{X}] \\ &= [bmatrix \mathcal{X}_1^1 \mathcal{X}_2^1 \mathcal{X}_3^1 \dots \mathcal{X}_{\mathcal{N}_{var}}^1; \dots \mathcal{X}_1^{\mathcal{N}_{pop}} \mathcal{X}_2^{\mathcal{N}_{pop}} \dots \mathcal{X}_{\mathcal{N}_{var}}^{\mathcal{N}_{pop}} bmatrix \end{aligned} \quad (10)$$

where $\mathcal{N}_{pop}$ and $\mathcal{N}_{var}$ are the initial population (number of raindrops) and the variation population, respectively. The values of the provided cost function ($\mathcal{C}$) are obtained from Equation (11).

$$\mathcal{C}_i = Cost_i = f(\mathcal{X}_1^i, \mathcal{X}_2^i, \dots, \mathcal{X}_{\mathcal{N}_{var}}^i), i = 1, 2, 3, \dots, \mathcal{N}_{pop} \quad (11)$$

where $\mathcal{C}_i$ is each drop's target value. The $\mathcal{N}_{SR}$ number of the best droplets is chosen as the sea and river after the $\mathcal{N}_{pop}$ number of raindrops (first stage) is formed. The raindrops that fall in the fewest numbers are referred to as the sea. $\mathcal{N}_{SR}$ is the total of the rivers (a parameter that is applied) and the sea (Equation (12)). The remaining population is calculated using Equation (13) and includes streams that may directly flow into rivers or the sea.

$$\mathcal{N}_{SR} = \text{Number of river} + \overset{sea}{1} \quad (12)$$

$$\mathcal{N}_{Raindrop} = \mathcal{N}_{pop} - \mathcal{N}_{SR} \quad (13)$$

To identify or assign raindrops to rivers and seas using Equation (14), based on the force of the flow.

$$\mathcal{N}_{Sn} = \text{round} \left\{ \left| \frac{Cost_n}{\sum_{i=1}^{\mathcal{N}_{SR}} Cost_i} \right| \times \mathcal{N}_{Raindrops} \right\}, n = 1, 2, \dots, \mathcal{N}_{SR} \quad (14)$$

where $\mathcal{N}_{Sn}$ is the total number of streams that empty into a given river or body of water. The randomly selected distance and a stream that flows up to the river are used in Equation (15) to establish a connection between them.

$$\mathcal{X} \in (0, \mathcal{C} \times d), \mathcal{C} > 1 \quad (15)$$

Using Equations (16) and (17), updated positions of streams and rivers can be found.

$$\mathcal{X}_{Stream}^{i+1} = \mathcal{X}_{Stream}^i + rand \times \mathcal{C} \times (\mathcal{X}_{River}^i - \mathcal{X}_{Stream}^i) \quad (16)$$

$$\mathcal{X}_{River}^{i+1} = \mathcal{X}_{River}^i + rand \times \mathcal{C} \times (\mathcal{X}_{Sea}^i - \mathcal{X}_{River}^i) \quad (17)$$

where $rand$ is a consistent random number between 0 and 1. The location of the river and the stream will change if the stream provides a better solution than the river to which it is connected. Rivers and seas are likewise susceptible to this transformation in the same way. One of the most crucial reasons impeding the algorithm's speedy convergence and becoming stuck in the local minimum is evaporation. When seawater enters rivers or streams, the evaporation process causes it to evaporate once more. Pseudo-code is used to show how to determine whether a river drains into the sea (Equation (18)).

$$if \left| \mathcal{X}_{Sea}^i - \mathcal{X}_{River}^i \right| \quad (18)$$

where d_{max} is a negligible number (close to zero). As a result, the river has reached the sea if the distance between it and it is smaller than d_{max} . The evaporation process is impacted in this condition, and precipitation will start once there has been enough evaporation. The search intensity close to the sea is controlled by d_{max} (optimal solution). Equation (19) reduces the value of max in each step.

$$d_{max}^{i+1} = d_{max}^i - \frac{d_{max}^i}{maxiteration} \quad (19)$$

After evaporation is accomplished, rainfall is applied. As rain falls, fresh raindrops form streams in various spots. The updated position of streams can be found using Equation (20).

$$\mathcal{X}_{Stream}^{new} = \mathcal{L}b + rand \times (\mathcal{U}b - \mathcal{L}b) \quad (20)$$

where $\mathcal{L}b$ and $\mathcal{U}b$ stand for the problem's upper and lower limits, respectively. The best newly created raindrops are thought of as rivers, and the remaining raindrops are thought of as fresh streams feeding the rivers. Equation (21) is applied to improve the computational efficiency and convergence rate of the approach for constrained problems.

$$\mathcal{X}_{Stream}^{new} = \mathcal{X}_{Sea} + \sqrt{\mu} \times randn(1, \mathcal{N}_{var}) \quad (21)$$

where the search area near the sea is represented by the factor. Small values cause the algorithm to search in a narrower area near the water, while large values increase the likelihood of leaving the possible area. Its proper value is set to 0.1 [50].

3.4. Hybrid Water Cycle-Moth-Flame Optimization Algorithm (WCAMFO)

Theoretically, the WCA and MFO algorithms showed that the water cycle (WCA) performed admirably throughout the discovery phase but poorly during operator-performed exploitation. As an alternative, the MFO exploited the space very well by using its spiral movement capability, although they frequently remained inherently unable to explore the solution space and was trapped in local optima [45,51]. As a result, by combining the benefits of the two methods, a hybrid algorithm can be a good alternative and help to further improve the optimization problem. Two of the major benefits of merging the WCA and the MFO, according to Khalil pourazari and Khalil pourazary [45] are that: the initial improvement was that the position in the WCA was updated using the spiral movement of the moths. The second enhancement is to improve the precipitation process. The so-called "Levy flight", from a mathematical perspective, allows the streams to arrange themselves better:

$$\mathcal{X}_{i+1} = \mathcal{X}_i + Levy(dim) \otimes \mathcal{X}_i \quad (22)$$

3.5. Performance Evaluation

Four statistical indicators of root mean squared error (RMSE), mean absolute error (MAE), determination coefficient (R^2) and Nash efficiency (NSE) were utilized for performance examination of RVFL, RVFL-WCA, RVFL-MFO, and RVFL-WCAMFO models for modeling air temperature. RMSE, MAE, NSE and R^2 can be expressed, respectively, by:

$$RMSE = \sqrt{\frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} [(T_c)_i - (T_p)_i]^2} \quad (23)$$

$$MAE = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} |(T_c)_i - (T_p)_i| \quad (24)$$

$$NSE = 1 - \frac{\sum_{i=1}^{\mathcal{N}} [(T_c)_i - (T_p)_i]^2}{\sum_{i=1}^{\mathcal{N}} [(T_c)_i - \overline{(T_c)}]^2} \quad (25)$$

$$R^2 = \frac{\sum_{i=1}^{\mathcal{N}} [(T_c)_i - \overline{(T_c)}][(T_p)_i - \overline{(T_p)}]}{\sum_{i=1}^{\mathcal{N}} [(T_p)_i - \overline{(T_p)}]^2 \sum_{i=1}^{\mathcal{N}} [(T_c)_i - \overline{(T_c)}]^2} \quad (26)$$

where T_c , T_p , $\overline{T_c}$ and $\overline{T_p}$ indicate calculated, predicted, mean calculated, and mean predicted air temperature, respectively, and $\mathcal{N}$ is the quantity of data.

4. Results and Discussion

In this section, the single optimization techniques (RVFL, WCA, MFO, WCAMFO, and all combinations) for modeling air temperature were technically discussed and comparatively analyzed. Prior to the computational schema, eight different input combinations were proposed for the development of each algorithm (M1, M2, and M3) (Table 1). M1, M2, and M3 testing datasets were formed on the basis of cross-validation technique. According to this technique, the whole dataset is divided into three equal parts and all three testing datasets are formed using equally 1/3 ratio of the whole dataset, whereas during analyzing one testing dataset, the remaining 2/3 of the whole dataset is used as a training dataset. It is worth noting that selecting input combinations for building machine learning is paramount and has been recommended in several works of literature [52–54]. Subsequently, the predictive models were assessed by four statistical matrices, including NSE, MAE, RMSE, and R^2 . For modeling nonlinear environmental variables such as air temperature, it is always recommended to employ several performance indicators to comprehend the stochastic influence of air temperature estimation based on the global and regional scale of climate change impact. According to Benaafi et al., [55] evaluation criteria should involve at least one error-of-fit (e.g., MAE) and one goodness-of-fit (e.g., NSE) for any reliable analysis.

Table 1. The input combinations used for model development.

Input Combinations	Variables
(i)	Pt
(ii)	Pt, Pt-1
(iii)	Pt, Pt-1, Pt-2
(iv)	Pt, Pt-1, Pt-2, Pt-3
(v)	Tt-1
(vi)	Tt-1, Tt-2
(vii)	Tt-1, Tt-2, Tt-3
(viii)	best p, best T

Furthermore, another essential part in AI-based models that receives less attention is controlling the hyper-tuning parameters in the modeling process. The parameters' optimization setting of each algorithm as tabulated in Table 2 were used for attaining the training and validation phase of each model's combination. The external configuration and the un-estimated parameters from each model are referred to as hyper parameter.

Table 2. Parameters setting of each optimization algorithm.

RVFL	Number of hidden neurons	200
	Activation function	radial basis
	Bias	1
	link between the input and output	1
	Updating method	Ridge regression
WCA	d_{max}	0.0001
	N_{sr}	8
MFO	constant to define the shape of the logarithmic spiral (b)	1
	l	$[-1, 1]$

Table 2. *Cont.*

WCAMFO	As in both WCA and MFO	
All Algorithms		
	Population	50
	Number of iterations	100
	Number of runs for each Algorithm	15

These parameters were tuned to determine the best values to solve a certain predictive model. It should be noted that different approaches were used for the determination of the best tuning parameters to find the most skillful predictions, for example, trial and error, rules of thumb, etc. Despite less attention than was previously observed in modeling the monthly and daily air temperature using machine learning, this study was the first work to employ the above-mentioned techniques to model air temperature to the best of the authors' knowledge.

4.1. Results for Monthly Air Temperature Modeling

The predictive outcomes in terms of several performance indicators for the RVFL monthly scale approach are presented in Table 3. The results indicated that M1 (viii) with the input combination of (best p, best T) attained the maximum values of goodness-of-fit of $R^2 = 0.921$, $NSE = 0.92$, and minimum values of $RMSE = 1.450$, $MAE = 1.203$ in comparison with M2 and M3 combinations. Furthermore, the best predictive combination of M2 and M3 were found to be M2 (viii), and M3 (viii), respectively. The outcomes of the model RVFL for the monthly time scale depicted that model combination (viii) with (best p, best T) emerging as the most reliable input with the highest accuracy in the testing phase. The fitting comparison of models shows that the best models M1 (viii) increased the predictive skills with giving range between 3% to 37% of the rest combination (i, ii, iii, iv, and vii), M2 (viii) within 4% to 42%, and lastly M3 (vii) within the range of 1% to 42%. The overall comparison of all the three models indicated that M1 (viii) with the smallest absolute error proved merit than the other models. This can be justified by considering the mean values in Table 3. In addition, Figure 4 shows the scatterplots of the observed and predicted temperature by different RVFL-based models in the test period using the best input combination for a monthly time scale.

Table 3. The results of the model RVFL for monthly time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R2	RMSE	MAE	NSE	R ²
M1	i	3.202	2.410	0.517	0.520	3.322	2.290	0.541	0.553
	ii	3.103	2.342	0.546	0.557	3.230	2.301	0.565	0.570
	iii	3.042	2.291	0.563	0.570	3.321	2.403	0.539	0.547
	iv	2.911	2.223	0.601	0.604	3.213	2.502	0.569	0.595
	v	2.844	2.252	0.617	0.619	2.852	2.230	0.632	0.632
	vi	1.922	1.632	0.824	0.826	1.794	1.481	0.866	0.867
	vii	1.740	1.467	0.857	0.859	1.591	1.294	0.885	0.886
	viii	1.401	1.124	0.911	0.912	1.450	1.203	0.912	0.921
	Mean	2.521	1.968	0.680	0.683	2.597	1.963	0.689	0.696

Table 3. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R2	RMSE	MAE	NSE	R ²
M2	i	3.291	2.532	0.523	0.525	3.591	2.362	0.385	0.432
	ii	3.153	2.323	0.563	0.564	3.172	2.433	0.521	0.539
	iii	3.160	2.361	0.562	0.572	3.533	2.302	0.405	0.463
	iv	2.804	2.144	0.655	0.656	3.301	2.583	0.479	0.567
	v	2.892	2.280	0.634	0.634	2.833	2.271	0.617	0.617
	vi	1.833	1.537	0.853	0.853	1.962	1.660	0.814	0.816
	vii	1.831	1.532	0.856	0.857	1.951	1.653	0.817	0.819
	viii	1.420	1.150	0.913	0.913	1.750	1.372	0.853	0.861
	Mean	2.548	1.982	0.695	0.697	2.762	2.080	0.611	0.639
M3	i	3.201	2.380	0.545	0.549	3.431	2.390	0.452	0.471
	ii	2.973	2.162	0.607	0.608	3.320	2.232	0.487	0.502
	iii	2.990	2.183	0.601	0.609	3.373	2.492	0.471	0.481
	iv	2.372	1.701	0.749	0.749	3.241	2.643	0.510	0.513
	v	2.751	2.162	0.663	0.665	2.972	2.401	0.621	0.621
	vi	1.872	1.560	0.844	0.844	1.880	1.603	0.836	0.836
	vii	1.473	1.173	0.904	0.910	1.661	1.350	0.882	0.888
	viii	1.440	1.140	0.902	0.905	1.494	1.227	0.896	0.897
	Mean	2.384	1.808	0.727	0.730	2.672	2.042	0.644	0.651
Best DS	Best IC, MN	1.370	1.082	0.917	0.918	1.414	1.161	0.918	0.923

In addition to the application of RVFL models, the hybrid optimization models using RVFL-WCA, RVFL-MFO, and RVFL-WCAMFO for monthly air temperature modeling were employed in this study and the estimated outcomes were generated in Tables 4–6.

Table 4. The results of the model RVFL-WCA for monthly time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R2	RMSE	MAE	NSE	R ²
M1	i	3.122	2.280	0.541	0.542	3.221	2.370	0.567	0.574
	ii	2.961	2.142	0.585	0.588	3.263	2.503	0.556	0.580
	iii	2.970	2.133	0.581	0.583	3.172	2.492	0.580	0.595
	iv	2.522	1.821	0.700	0.701	2.920	2.191	0.646	0.653
	v	2.803	2.190	0.630	0.632	2.933	2.290	0.642	0.646
	vi	1.810	1.521	0.846	0.853	1.702	1.401	0.880	0.895
	vii	1.492	1.242	0.895	0.897	1.460	1.170	0.911	0.914
	viii	1.153	0.960	0.931	0.932	1.352	1.021	0.919	0.930
	Mean	2.354	1.786	0.714	0.716	2.503	1.930	0.713	0.723

Table 4. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M2	i	3.212	2.351	0.548	0.549	3.341	2.581	0.468	0.469
	ii	3.051	2.150	0.590	0.590	3.170	2.222	0.521	0.541
	iii	2.960	2.032	0.614	0.614	3.352	2.661	0.464	0.492
	iv	2.672	1.983	0.687	0.687	3.103	1.940	0.540	0.576
	v	2.823	2.191	0.651	0.652	2.741	2.153	0.641	0.641
	vi	1.661	1.350	0.878	0.885	1.770	1.502	0.851	0.855
	vii	1.303	1.042	0.926	0.931	1.522	1.181	0.890	0.901
	viii	1.280	0.853	0.949	0.949	1.491	1.010	0.894	0.903
	Mean	2.370	1.744	0.730	0.732	2.561	1.906	0.659	0.672
M3	i	3.042	2.181	0.589	0.590	3.311	2.272	0.490	0.503
	ii	2.801	1.953	0.651	0.651	3.203	2.240	0.522	0.527
	iii	2.750	2.021	0.662	0.663	3.241	2.252	0.512	0.522
	iv	2.182	1.602	0.788	0.788	3.002	2.133	0.581	0.593
	v	2.521	1.941	0.718	0.718	2.810	2.202	0.631	0.631
	vi	1.603	1.313	0.886	0.886	1.611	1.341	0.879	0.879
	vii	1.242	0.981	0.932	0.932	1.422	1.080	0.917	0.920
	viii	1.210	0.890	0.941	0.941	1.403	1.152	0.909	0.909
	Mean	2.169	1.610	0.771	0.771	2.500	1.834	0.680	0.686
Best DS	Best IC, MN	1.092	0.944	0.933	0.937	1.162	0.871	0.944	0.946

Table 5. The results of the model RVFL-MFO for monthly time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M1	i	3.063	2.202	0.554	0.558	3.192	2.391	0.576	0.584
	ii	2.914	2.003	0.600	0.602	3.111	2.322	0.597	0.608
	iii	2.910	2.140	0.601	0.602	3.203	2.482	0.574	0.599
	iv	2.403	1.671	0.725	0.729	2.710	2.103	0.693	0.704
	v	2.561	1.952	0.687	0.690	2.702	2.060	0.697	0.697
	vi	1.362	1.033	0.911	0.913	1.253	1.002	0.935	0.936
	vii	1.140	0.891	0.938	0.939	1.220	0.991	0.938	0.939
	viii	0.972	0.853	0.977	0.977	1.051	0.790	0.954	0.956
	Mean	2.166	1.593	0.749	0.751	2.305	1.768	0.746	0.753

Table 5. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M2	i	3.171	2.313	0.559	0.559	3.191	2.360	0.515	0.534
	ii	3.040	2.122	0.590	0.594	3.020	2.202	0.565	0.569
	iii	2.873	2.030	0.635	0.637	3.261	2.353	0.492	0.545
	iv	2.492	1.771	0.725	0.727	2.810	1.902	0.624	0.636
	v	2.331	1.662	0.760	0.762	2.452	1.711	0.713	0.715
	vi	1.241	0.960	0.933	0.934	1.403	1.040	0.907	0.910
	vii	1.152	0.902	0.941	0.942	1.291	1.002	0.921	0.923
	viii	0.910	0.863	0.968	0.971	1.191	0.912	0.932	0.938
	Mean	2.151	1.578	0.764	0.766	2.327	1.685	0.709	0.721
M3	i	3.012	2.091	0.597	0.597	3.261	2.281	0.505	0.517
	ii	2.863	1.942	0.636	0.636	3.212	2.512	0.521	0.540
	iii	2.841	1.991	0.642	0.642	3.280	2.211	0.499	0.512
	iv	2.050	1.480	0.813	0.813	2.951	1.933	0.596	0.626
	v	2.452	1.903	0.733	0.733	2.733	1.991	0.653	0.654
	vi	1.193	0.911	0.937	0.938	1.422	1.170	0.907	0.907
	vii	1.021	0.812	0.954	0.954	1.250	0.962	0.928	0.928
	viii	0.965	0.784	0.971	0.973	1.331	1.110	0.917	0.921
	Mean	2.050	1.489	0.785	0.786	2.430	1.771	0.691	0.701
Best DS	Best IC, MN	0.633	0.474	0.980	0.981	1.072	0.833	0.952	0.954

Table 6. The results of the model RVFL-WCAMFO for monthly time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M1	i	2.930	2.062	0.592	0.594	3.132	2.290	0.593	0.603
	ii	2.771	1.880	0.636	0.639	3.091	2.213	0.602	0.614
	iii	2.674	1.892	0.661	0.663	3.110	2.272	0.596	0.608
	iv	2.240	1.561	0.760	0.763	2.684	2.050	0.702	0.718
	v	2.282	1.600	0.751	0.754	2.261	1.583	0.787	0.790
	vi	0.983	0.724	0.954	0.955	1.172	0.990	0.950	0.953
	vii	0.910	0.691	0.960	0.961	1.160	0.933	0.944	0.946
	viii	0.862	0.643	0.982	0.984	0.962	0.812	0.962	0.963
	Mean	1.957	1.382	0.787	0.789	2.197	1.643	0.767	0.774

Table 6. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M2	i	3.091	2.152	0.580	0.580	2.952	2.070	0.585	0.586
	ii	2.960	2.023	0.612	0.614	2.811	2.012	0.624	0.624
	iii	2.792	1.960	0.654	0.658	3.003	2.090	0.570	0.579
	iv	2.310	1.613	0.765	0.765	2.642	1.941	0.667	0.686
	v	2.242	1.551	0.777	0.780	2.383	1.591	0.729	0.732
	vi	0.911	0.703	0.962	0.964	1.112	0.872	0.941	0.944
	vii	0.884	0.685	0.963	0.966	1.070	0.865	0.946	0.949
	viii	0.860	0.623	0.984	0.986	1.037	0.860	0.939	0.940
	Mean	2.006	1.414	0.787	0.789	2.126	1.538	0.750	0.755
M3	i	2.890	2.021	0.628	0.628	3.192	2.410	0.525	0.533
	ii	2.692	1.862	0.677	0.677	3.150	2.101	0.539	0.550
	iii	2.653	1.870	0.687	0.687	3.173	2.194	0.531	0.540
	iv	1.841	1.323	0.849	0.849	2.912	1.942	0.605	0.621
	v	2.110	1.461	0.802	0.802	2.591	1.703	0.688	0.695
	vi	0.992	0.774	0.956	0.956	1.214	0.872	0.932	0.932
	vii	0.903	0.711	0.964	0.964	1.122	0.880	0.941	0.942
	viii	0.880	0.635	0.985	0.985	1.223	0.921	0.930	0.931
	Mean	1.870	1.332	0.819	0.819	2.322	1.628	0.711	0.718
Best DS	Best IC, MN	0.531	0.383	0.983	0.986	0.824	0.661	0.972	0.973

The concept of coupling RVFL with three different optimization algorithms was attributed to its universal approximation ability and fast training speed as reported by several literatures. Hence, during the monthly air temperature modeling using RVFL-WCA, the M1 (viii) with parameters combination (Tt-1, Tt-2, Tt-3) in Table 4 supported the best against all the combination (M1, and M2). The quantitative comparison of the best outcomes with regards to absolute error of the M1, M2, and M3 demonstrated that M1 (viii) reduced the prediction error by approximately 12% and 6% for M2, and M3, respectively. This outcomes was in line with the work conducted by Smith et al., [56] that improved the estimation accuracy of air temperature using ANN models and obtained the hourly prediction accuracy more than 90% in terms of fitting of the models (see Figure 5). The major difference between our work and that of Smith et al. [56] is the employment of recently developed state-of-the-art models supported by optimization algorithms that provide desirable results with fewer input combinations. The overall results can also be depicted using violin plots as presented in Figure 6.

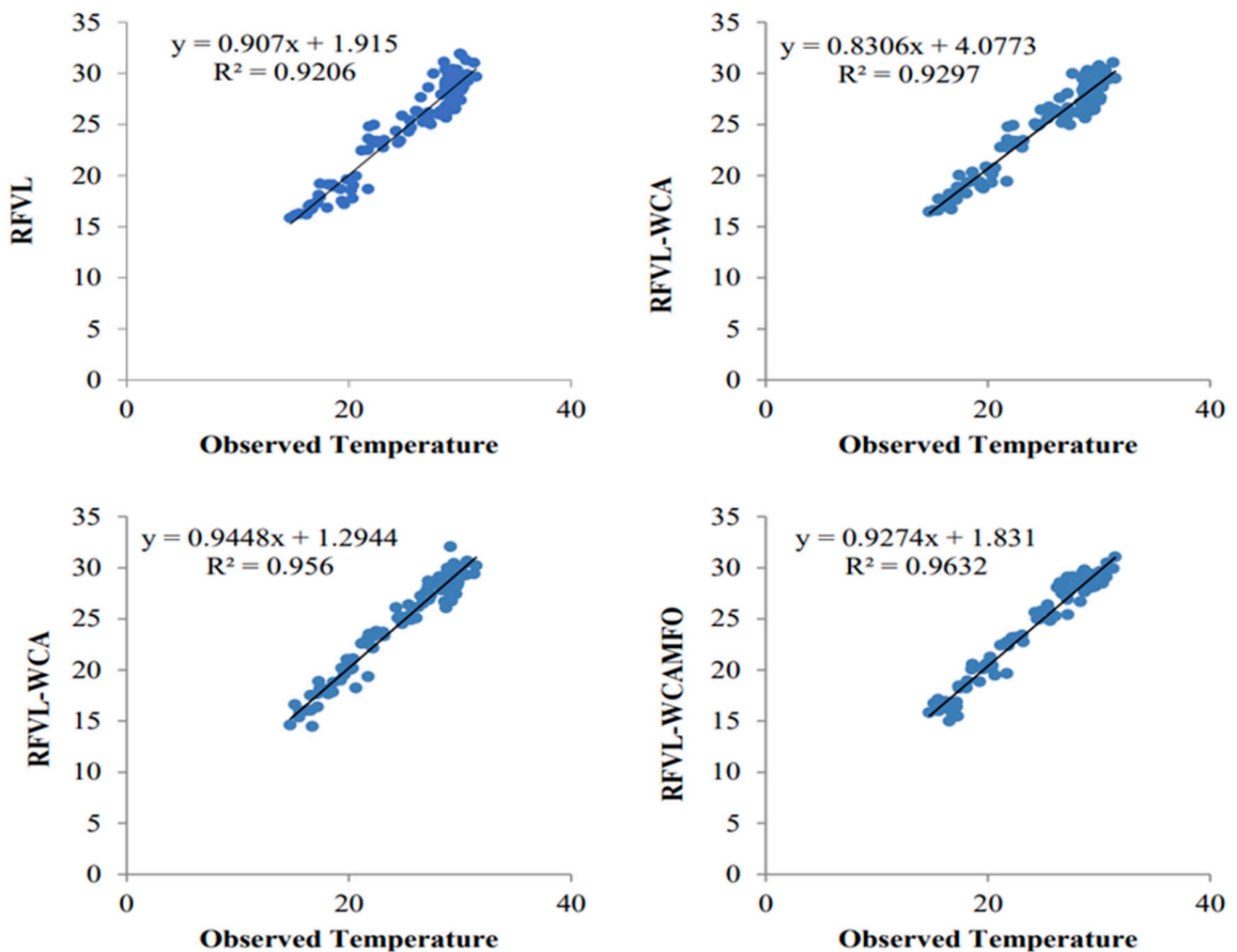


Figure 4. Scatterplots of the observed and predicted Temperature by different RVFL-based models in the test period using best input combination for monthly time scale.

Based on the present outcomes, it is worth mentioning that computational approaches play an essential role in handling any type of chaotic pattern. The inputs and response variables justified merit with reasonable accuracy using our established models. To give credit and compare our established outcomes, this study was compared with the selected state-of-the-art models for examples [57–63] in terms of the popularity of the ML model and proved satisfactory. Other MLS models could be simulated similarly using the same problems.

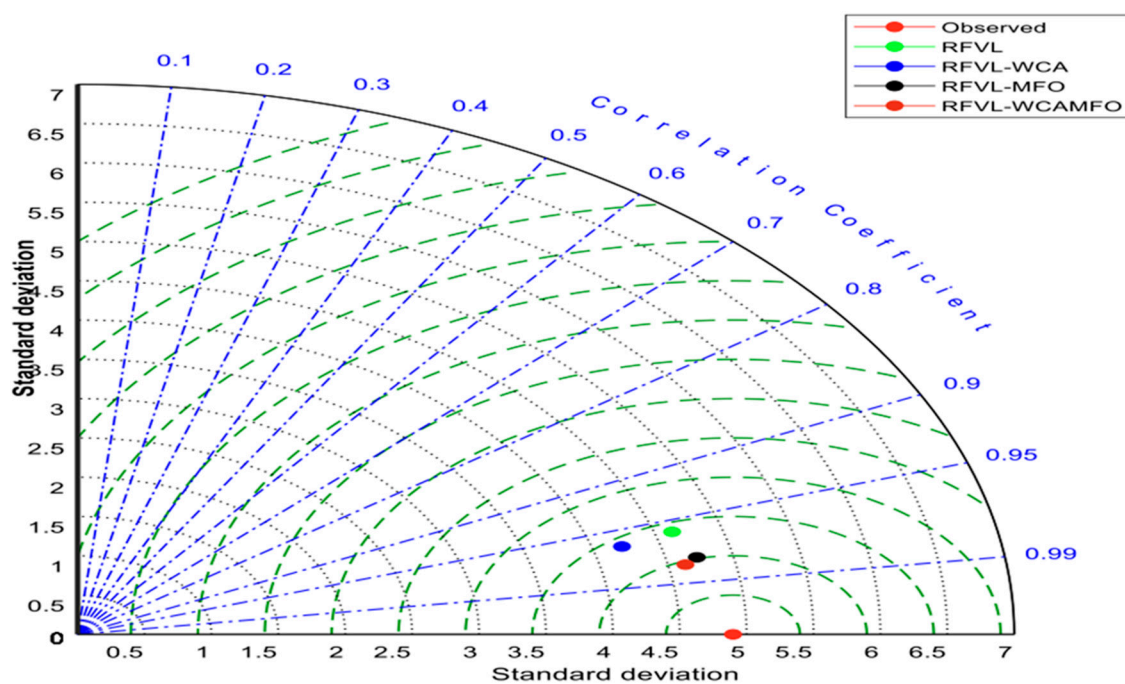


Figure 5. Taylor diagrams of the estimated temperature by different RVFL-based models in the test period for monthly time scale.

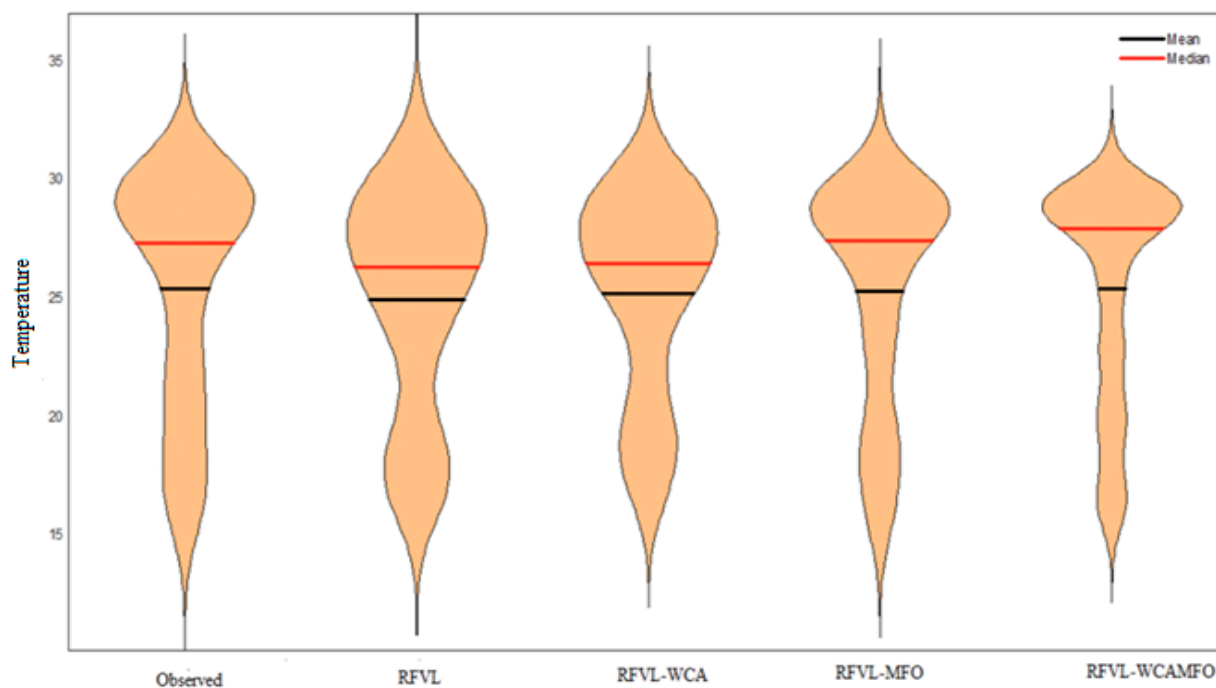


Figure 6. Violin Charts estimated temperature by different RVFL-based models in the test period for monthly time scale.

4.2. Results for Daily Air Temperature Modeling

In this section, the daily air temperature was also simulated at study locations. Note that the combination of modeling the short and long extreme weather changes for air temperature, for example, daily and monthly modeling received less, or no attention based on the authors' knowledge, despite other hydrological time series processes such as rainfall and run-off already reporting similar scenarios [57,58]. Air temperature modeling and prediction of air pollution serves as major environmental monitoring in different regions. This study is devoted to demonstrating daily and monthly information to come up with satisfactory estimation methodologies for the decision-maker and concern authorities with regards to short- and long-term environmental sustainability. The simulated results for a daily time scale of air temperature for RVFL, RVFL-WCA, RVFL-MFO, and RVFL-WCAMFO are presented in Table 7, Table 8, Table 9, and Table 10, respectively. The results of the daily time scale of air temperature in Figure 7 demonstrated that the first four combinations produce poor to marginal accuracy by comparing the evaluation matrices (MAE, NSE, R^2 , and RMSE). However, during the daily scale, the models with the combination M1-M3 (vii) displayed high accuracy against all the combinations. For example the assessment matrices, the best model, demonstrated that M1 (vii) = MAE (0.947), NSE (0.920), M2 (vii) = MAE (0.988), NSE (0.918), and M3 (vii) = MAE (0.960), NSE (0.92) for air temperature modeling. The detailed results for all the combinations are presented in terms of performance evaluation criteria in Table 7. It can be observed that single standalone models, i.e., RVFL, could produce reasonable accuracy for modeling air temperature. The feasibility of the hybrid state-of-the-art models for modeling air temperature were also presented using the hybrid optimization approach. Furthermore, other different AI models could also be used in the same manner for proper tuning of air temperature prediction. Table 8 demonstrates the results of the model RVFL-WCA for a daily time scale.

Table 7. The results of the model RVFL for daily time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R^2	RMSE	MAE	NSE	R^2
M1	i	5.385	4.471	0.026	0.031	5.614	4.325	0.028	0.033
	ii	5.374	4.445	0.028	0.033	5.528	4.536	0.031	0.036
	iii	5.262	4.389	0.034	0.038	5.416	4.427	0.042	0.047
	iv	5.113	4.301	0.039	0.043	5.289	4.276	0.048	0.055
	v	1.263	0.962	0.911	0.916	1.253	0.958	0.914	0.918
	vi	1.259	0.957	0.914	0.918	1.248	0.952	0.919	0.922
	vii	1.248	0.949	0.921	0.924	1.239	0.947	0.920	0.922
	viii	1.271	0.973	0.902	0.905	1.262	0.965	0.904	0.907
	Mean	3.272	2.681	0.472	0.476	3.356	2.673	0.476	0.481
M2	i	5.275	4.449	0.042	0.045	5.156	4.387	0.045	0.048
	ii	5.229	4.331	0.054	0.057	5.127	4.218	0.059	0.060
	iii	5.144	4.274	0.058	0.060	5.056	4.149	0.062	0.065
	iv	5.082	4.104	0.064	0.068	5.021	4.076	0.068	0.070
	v	1.271	0.982	0.904	0.909	1.316	1.016	0.903	0.905
	vi	1.266	0.974	0.911	0.915	1.302	1.003	0.908	0.910
	vii	1.255	0.958	0.924	0.926	1.296	0.988	0.915	0.918
	viii	1.280	0.987	0.897	0.900	1.324	1.022	0.895	0.898
	Mean	3.225	2.632	0.482	0.485	3.200	2.607	0.482	0.484

Table 7. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M3	i	5.278	4.440	0.043	0.045	5.126	4.327	0.038	0.040
	ii	5.258	4.426	0.046	0.049	5.107	4.308	0.041	0.042
	iii	5.167	4.381	0.050	0.052	5.087	4.268	0.045	0.048
	iv	5.105	4.265	0.055	0.058	5.014	4.192	0.051	0.053
	v	1.251	0.959	0.908	0.910	1.281	0.975	0.905	0.906
	vi	1.246	0.948	0.916	0.918	1.276	0.967	0.909	0.912
	vii	1.238	0.940	0.927	0.930	1.266	0.960	0.915	0.920
	viii	1.262	0.964	0.901	0.903	1.292	0.986	0.898	0.900
	Mean	3.226	2.665	0.481	0.483	3.181	2.623	0.475	0.478

Table 8. The results of the model RVFL-WCA for daily time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M1	i	4.835	4.070	0.038	0.042	5.131	4.337	0.037	0.039
	ii	4.721	3.918	0.083	0.085	4.997	4.154	0.086	0.090
	iii	4.642	3.828	0.113	0.113	4.904	4.049	0.120	0.125
	iv	4.587	3.772	0.134	0.134	4.844	3.991	0.142	0.147
	v	1.248	0.946	0.926	0.929	1.237	0.944	0.925	0.928
	vi	1.243	0.942	0.931	0.933	1.227	0.942	0.929	0.932
	vii	1.237	0.937	0.935	0.938	1.218	0.940	0.935	0.936
	viii	1.257	0.958	0.912	0.915	1.246	0.953	0.915	0.918
	Mean	2.971	2.421	0.509	0.511	3.101	2.539	0.514	0.518
M2	i	4.989	4.188	0.066	0.069	4.799	4.028	0.063	0.065
	ii	4.905	4.089	0.070	0.073	4.718	3.944	0.072	0.075
	iii	4.746	3.882	0.131	0.132	4.577	3.763	0.128	0.130
	iv	4.623	3.742	0.173	0.176	4.452	3.610	0.172	0.174
	v	1.253	0.951	0.928	0.931	1.294	1.005	0.917	0.920
	vi	1.247	0.944	0.932	0.934	1.276	0.969	0.932	0.932
	vii	1.233	0.938	0.935	0.937	1.227	0.931	0.937	0.937
	viii	1.258	0.958	0.918	0.920	1.301	1.011	0.912	0.914
	Mean	3.032	2.462	0.519	0.522	2.956	2.408	0.517	0.518

Table 8. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M3	i	4.982	4.187	0.063	0.066	4.840	4.056	0.061	0.065
	ii	4.826	3.967	0.096	0.098	4.714	3.863	0.088	0.085
	iii	4.786	3.950	0.110	0.113	4.674	3.852	0.104	0.100
	iv	4.708	3.868	0.140	0.141	4.607	3.779	0.130	0.126
	v	1.239	0.942	0.930	0.932	1.274	0.968	0.918	0.920
	vi	1.230	0.934	0.935	0.936	1.268	0.959	0.923	0.925
	vii	1.217	0.931	0.937	0.939	1.261	0.952	0.926	0.927
	viii	1.248	0.995	0.923	0.925	1.281	0.975	0.905	0.908
	Mean	3.030	2.472	0.517	0.519	2.990	2.426	0.507	0.507

Table 9. The results of the model RVFL-MFO for daily time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M1	i	4.782	3.921	0.070	0.070	5.048	4.138	0.072	0.072
	ii	4.601	3.671	0.131	0.131	4.825	3.875	0.143	0.148
	iii	4.485	3.519	0.174	0.174	4.680	3.699	0.197	0.194
	iv	4.383	3.386	0.211	0.211	4.544	3.536	0.244	0.240
	v	1.243	0.938	0.934	0.934	1.231	0.937	0.929	0.933
	vi	1.238	0.934	0.937	0.939	1.228	0.934	0.932	0.936
	vii	1.234	0.932	0.940	0.943	1.225	0.930	0.937	0.945
	viii	1.249	0.944	0.928	0.931	1.238	0.944	0.928	0.932
	Mean	2.902	2.281	0.541	0.542	3.002	2.374	0.551	0.553
M2	i	4.914	4.048	0.068	0.071	4.724	3.858	0.073	0.075
	ii	4.737	3.759	0.132	0.135	4.554	3.641	0.141	0.141
	iii	4.604	3.581	0.180	0.183	4.414	3.449	0.190	0.190
	iv	4.505	3.463	0.215	0.218	4.326	3.348	0.223	0.222
	v	1.240	0.941	0.931	0.933	1.220	0.930	0.928	0.930
	vi	1.236	0.939	0.934	0.936	1.218	0.926	0.935	0.938
	vii	1.230	0.935	0.937	0.941	1.214	0.923	0.939	0.940
	viii	1.247	0.935	0.925	0.928	1.228	0.936	0.922	0.925
	Mean	2.964	2.325	0.540	0.543	2.862	2.251	0.544	0.545

Table 9. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M3	i	4.926	4.076	0.071	0.073	4.760	3.857	0.067	0.069
	ii	4.696	3.720	0.142	0.146	4.632	3.647	0.117	0.121
	iii	4.546	3.522	0.199	0.202	4.540	3.492	0.151	0.157
	iv	4.413	3.348	0.244	0.246	4.476	3.401	0.175	0.184
	v	1.222	0.932	0.935	0.937	1.259	0.946	0.925	0.928
	vi	1.220	0.930	0.938	0.940	1.257	0.944	0.930	0.931
	vii	1.212	0.927	0.941	0.943	1.249	0.940	0.932	0.935
	viii	1.231	0.942	0.928	0.930	1.272	0.958	0.918	0.921
	Mean	2.933	2.300	0.550	0.552	2.931	2.273	0.527	0.531

Table 10. The results of the model RVFL-WCAMFO for daily time scale.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M1	i	4.751	3.887	0.074	0.074	5.012	4.132	0.076	0.076
	ii	4.596	3.657	0.133	0.133	4.819	3.865	0.149	0.146
	iii	4.451	3.462	0.187	0.187	4.672	3.690	0.203	0.197
	iv	4.330	3.303	0.230	0.230	4.537	3.532	0.249	0.243
	v	1.236	0.934	0.937	0.939	1.228	0.934	0.940	0.941
	vi	1.219	0.926	0.944	0.947	1.222	0.928	0.942	0.945
	vii	1.214	0.923	0.951	0.953	1.218	0.921	0.947	0.948
	viii	1.242	0.943	0.932	0.935	1.235	0.938	0.933	0.938
	Mean	2.880	2.254	0.549	0.550	2.993	2.368	0.557	0.556
M2	i	4.876	3.997	0.072	0.074	4.712	3.850	0.077	0.077
	ii	4.730	3.747	0.134	0.137	4.547	3.620	0.138	0.140
	iii	4.585	3.548	0.187	0.190	4.408	3.441	0.190	0.192
	iv	4.441	3.374	0.238	0.240	4.304	3.340	0.225	0.230
	v	1.237	0.938	0.935	0.940	1.222	0.927	0.935	0.937
	vi	1.169	0.922	0.944	0.947	1.212	0.919	0.938	0.941
	vii	1.094	0.917	0.951	0.953	1.204	0.911	0.940	0.945
	viii	1.248	0.946	0.930	0.932	1.229	0.935	0.928	0.930
	Mean	2.923	2.299	0.549	0.552	2.855	2.243	0.546	0.549

Table 10. Cont.

Dataset	Input Combination	Training				Testing			
		RMSE	MAE	NSE	R ²	RMSE	MAE	NSE	R ²
M3	i	4.878	3.978	0.075	0.078	4.749	3.848	0.069	0.073
	ii	4.693	3.708	0.145	0.147	4.629	3.643	0.119	0.123
	iii	4.527	3.481	0.203	0.206	4.536	3.507	0.155	0.159
	iv	4.382	3.306	0.254	0.256	4.471	3.406	0.181	0.187
	v	1.216	0.926	0.941	0.942	1.258	0.945	0.930	0.934
	vi	1.205	0.912	0.947	0.949	1.245	0.933	0.935	0.936
	vii	1.191	0.903	0.954	0.957	1.244	0.931	0.937	0.939
	viii	1.228	0.935	0.933	0.936	1.265	0.953	0.924	0.925
	Mean	2.915	2.269	0.557	0.559	2.925	2.271	0.531	0.535

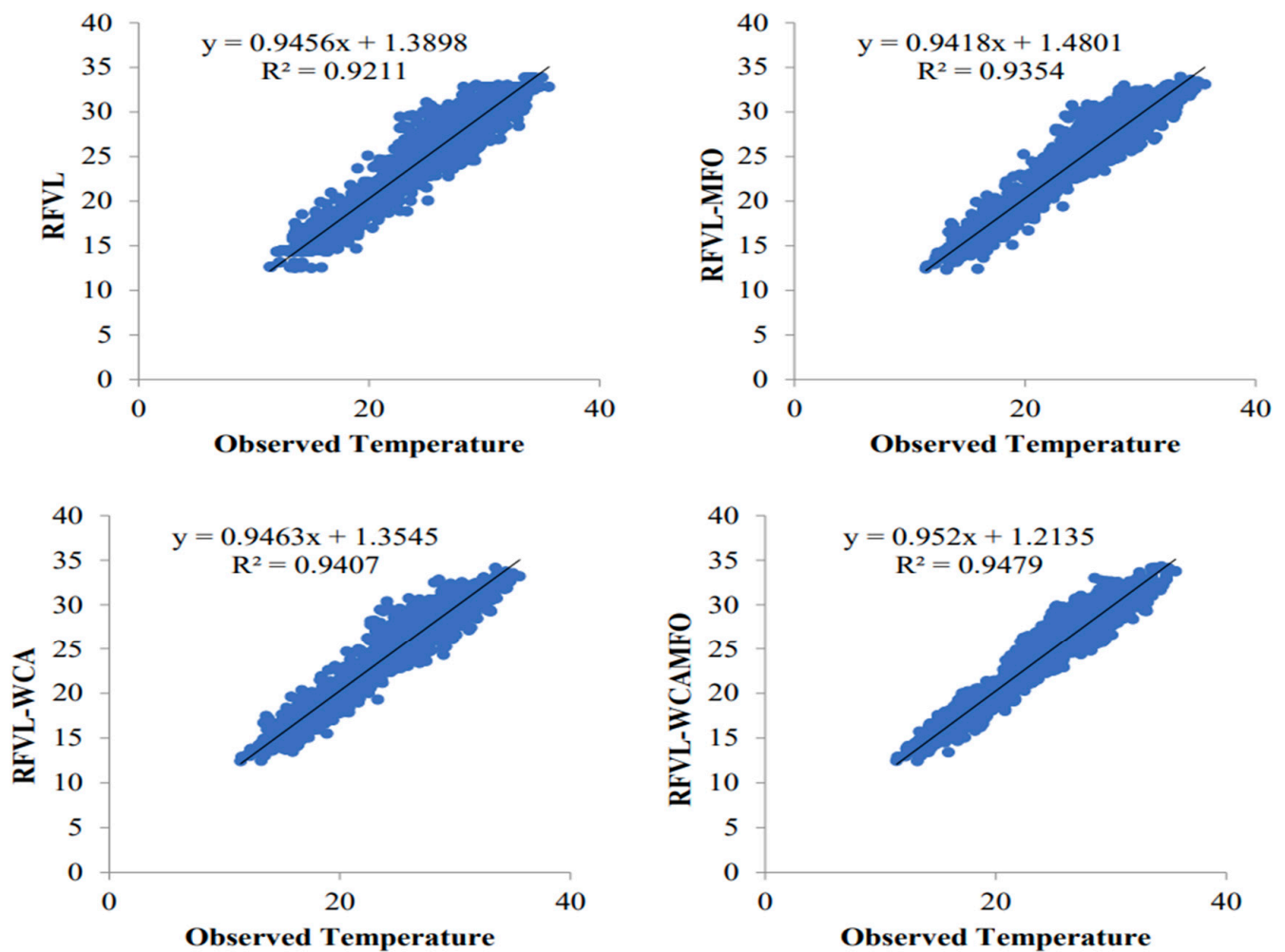


Figure 7. Scatterplots of the observed and predicted temperature by different RVFL-based models in the test period using best input combination for daily time scale.

It is worth mentioning that optimization algorithms were reported generally to optimize the tuning parameters and enhance the predictability of the models [59,60]. This was justified in the presented study where the combination (vii) became the most reliable among the other seven combinations. The numerical outcomes of RVFL-WCA for a daily time scale show that $M1(vii) = RMSE (1.218)$, $R^2 (0.936)$, $M2(vii) = RMSE (1.227)$, $R^2 (0.937)$, and $M3(vii) = RMSE (1.261)$, $R^2 (0.927)$. The capability of daily air temperature modeling is presented graphically in the scatter plot (Figure 7). The scatter plots of the observed and predicted air temperature by different RVFL-based models in the test period using the best input combination for a monthly time scale indicated that RVFL-WCAMFO attained the highest accuracy of more than 94.8% against the other, which have the accuracy ranging from 92–94%.

Similarly, the output of Table 9 for the results of the model RVFL-MFO for daily time scale indicated that $M1(vii) = MAE (0.930)$, $NSE (0.937)$, $M2(vii) = MAE (0.926)$, $NSE (0.939)$, $M3(vii) = MAE (0.940)$, $NSE (0.932)$.

For computational modeling of air temperature, it is obvious that step-ahead combination of temperature produces the best outcomes; this is in line with the presented study where the combination M1-M3 (vii) Tt-1, Tt-2, Tt-3 produced the best predictive skills. The obtained modeling architecture justified that minimum input parameters based on the accurate feature selection could lead to more improved prediction results than large combination. Similarly, the results show that all combinations with precipitation as inputs were less accurate, whereas temperature combination as previous inputs provided more accurate results. The predicted results of daily air temperature can also be visualized in the Taylor diagram using different matrices, as shown in Figure 8. The major advantage of the Taylor diagram is centered towards understanding the concept of correlation, standard deviation, and RMSE.

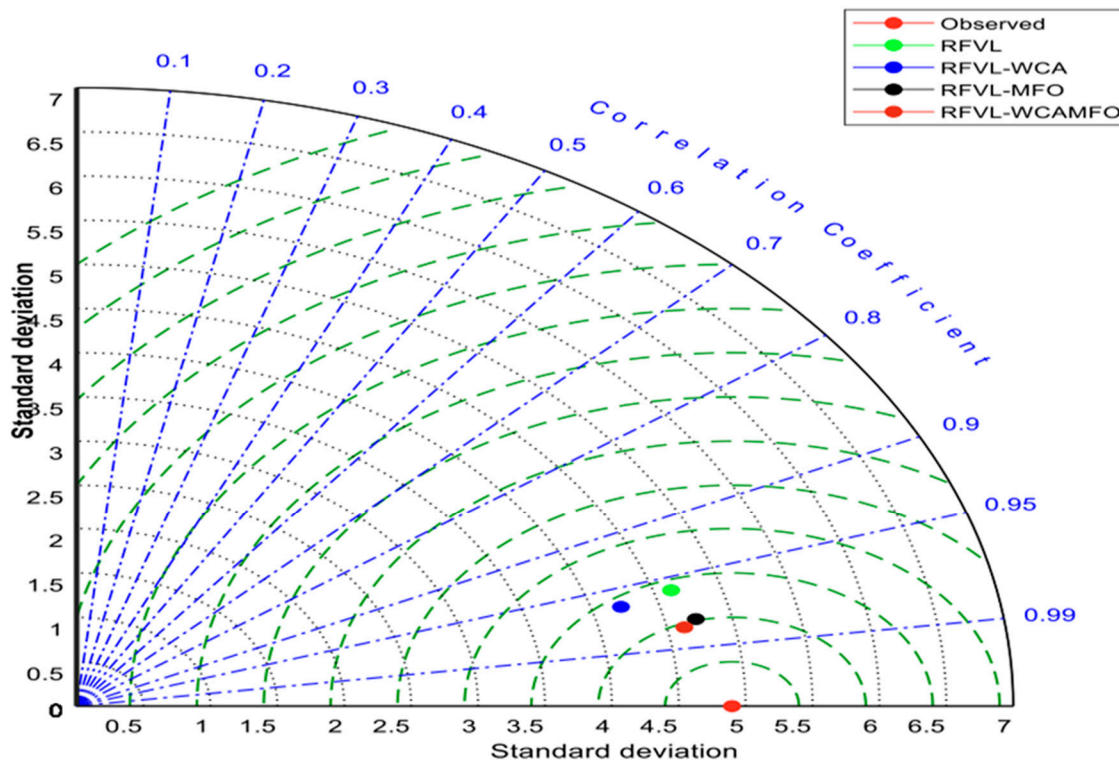


Figure 8. Taylor diagrams of the estimated Temperature by different RVFL-based models in the test period for daily time scale.

It can also be proved in this study demonstration that precipitation inputs have less effect on temperature, whereas what we see in literature, temperature inputs have good affect in modeling runoff and rainfall, for example, [61–64]. These results also demonstrated the necessity and importance of accurate temperature modeling as extreme rainfall events such as typhoon, cloud bursting, heat wave, and extreme runoff events such as drought and flood can be avoided if we can model air temperature precisely because this input has a very good affecting relationship with precipitation and runoff. It is a well-known fact that remarkable progress has been recorded in temperature modeling, but still some limitations exist, especially with a traditional theoretical approach. Table 10 presents the results of the model RVFL-WCAMFO for a daily time scale.

From the results it can be observed that the best models are still a temperature-related combination (vii). The summary of the numerical results for the best models demonstrated that M1(vii) = RMSE (1.218), MAE(0.921), NSE(0.947), R²(0.948), M2 (vii) = RMSE (1.204), MAE (0.911), NSE (0.940), R² (0.945), and M3 (vii) = RMSE (1.244), MAE (0.931), NSE (0.937), R² (0.939). Addition visualization based on the predicted results is presented in Figure 9.

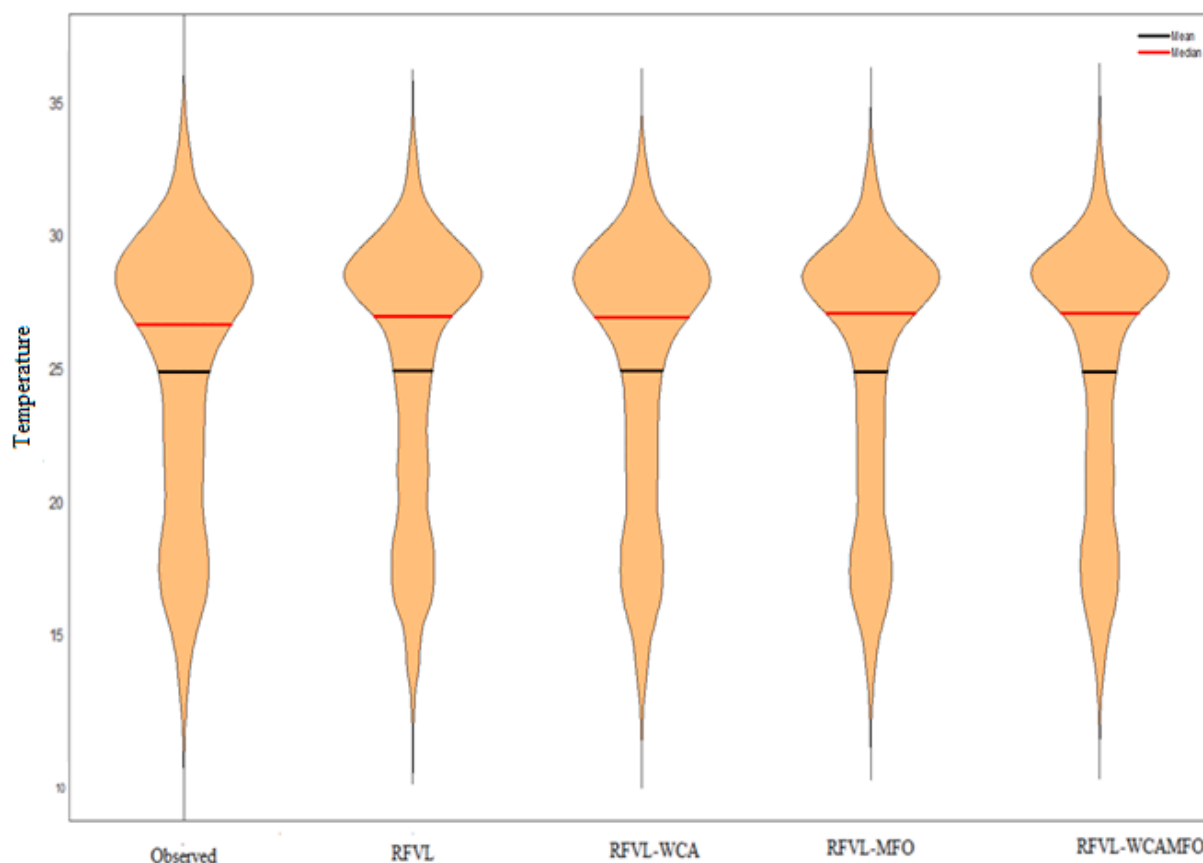


Figure 9. Violin Charts estimated temperature by different RVFL-based models in the test period for daily time scale.

5. Conclusions

The prediction accuracy of RVFL with three hybrid metaheuristic models, namely RVFL-WCA, RVFL-MFO, and RVFL-WCAMFO, for predicting the air temperature of the Rajshahi station of western Bangladesh was assessed in this paper. The standalone RVFL and three metaheuristic models were examined using three algorithms, including M1, M2, and M3, with eight different input combinations with the help of RMSE, MAE, and R² performance evaluation indexes, including scatter plots, and Taylor and violin charts. The metaheuristic algorithms made the single RVFL more accurate when it was used to simulate (in the training stage) and predict (in the testing stage) the monthly air temperature

based on M1, M2, and M3. The monthly air temperature modeling using a standalone RVFL model showed that the best model, M1 (viii), increased the predictive skills by giving variations from 3% to 37% of the remaining combinations (i, ii, iii, iv, and vii), M2 (viii) within 4% to 42%, and lastly, M3 (vii) within the range of 1% to 42%. The monthly air temperature modeling using hybrid optimization RVFL-WCA revealed that the M1 (viii) with parameter combinations (Tt-1, Tt-2, Tt-3) indicated the best against all the combinations (M1 and M2). The RMSE of the M1, M2, and M3 demonstrated that M1 (viii) reduced the prediction error by approximately 12% and 6% for M2 and M3, respectively, for the monthly time scale. During the daily scale, the RVFL-WCAMFO models with the combination (vii) displayed high accuracy against all the combinations for air temperature modeling. They have attained more than 94.8% accuracy against the other three models, with an accuracy range of 92–94%, though single standalone models, i.e., RVFL, could produce reasonable accuracy for modeling air temperature. The precision comparison of the algorithms exposed that the precision positions were in descending order: RVFL-WCAMFO > RVFL-WCA > RVFL-MFO in forecasting air temperature. According to the outcomes of this research, the use of hybrid metaheuristic i.e., RVFL-WCAMFO is suggested for air temperature prediction. Thus, the suggested model can be helpful for policymakers in alleviating the effects of temperature and recommending effective plans for agricultural crop production. The key limitation of this research was the use of input data from just one station in western Bangladesh to examine the accuracy of the models. In future studies, these models can be evaluated using more datasets from different regions. These suggested advanced hybrid models also can be evaluated with other different hybrid machine learning models in the future studies.

Author Contributions: Conceptualization: R.M.A., S.G.M., F.A. and Z.C.; formal analysis: R.R.M. and A.R.M.T.I.; validation: R.M.A., F.A., Z.C., S.G.M. and A.R.M.T.I.; supervision: A.R.M.T.I. and Z.C.; writing—original draft: R.M.A., S.I.A., R.R.M., S.G.M., Z.C., F.A. and A.R.M.T.I.; visualization: R.M.A. and S.I.A.; investigation: S.I.A., R.R.M. and F.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study will be available on an interesting request from the corresponding author.

Conflicts of Interest: There is no conflict of interest in this study.

References

1. Allan, R.P.; Hawkins, E.; Bellouin, N.; Collins, B. *Summary for Policymakers*; IPCC: Geneva, Switzerland, 2021.
2. Nita, I.A.; Sfică, L.; Voiculescu, M.; Birsan, M.V.; Micheu, M.M. Changes in the global mean air temperature over land since 1980. *Atmos. Res.* **2022**, *279*, 106392. [[CrossRef](#)]
3. Alvi, U.; Suomi, J.; Käyhkö, J. A cost-effective method for producing spatially continuous high-resolution air temperature information in urban environments. *Urban Clim.* **2022**, *42*, 101123. [[CrossRef](#)]
4. Yu, M.; Xu, F.; Hu, W.; Sun, J.; Cervone, G. Using Long Short-Term Memory (LSTM) and Internet of Things (IoT) for localized surface temperature forecasting in an urban environment. *IEEE Access* **2021**, *9*, 137406–137418. [[CrossRef](#)]
5. Cho, H.J.; Cheon, S.Y.; Jeong, J.W. Preliminary study on air-to-air latent heat exchanger fabricated using hollow fiber composite membrane for air-conditioning applications. *Energy Convers. Manag.* **2022**, *251*, 115000. [[CrossRef](#)]
6. Lembrechts, J.J.; Hoogen, J.V.D.; Aalto, J.; Ashcroft, M.B.; De Frenne, P.; Kemppinen, J.; Kopecký, M.; Luoto, M.; Maclean, I.M.D.; Crowther, T.W.; et al. Global maps of soil temperature. *Glob. Change Biol.* **2022**, *28*, 3110–3144. [[CrossRef](#)]
7. Di Blasi, C.; Renzi, M.; Michelozzi, P.; Donato, F.D.; Scortichini, M.; Davoli, M.; Forastiere, F.; Mannucci, P.M.; Stafoggia, M. Association between air temperature, air pollution and hospital admissions for pulmonary embolism and venous thrombosis in Italy. *Eur. J. Intern. Med.* **2022**, *96*, 74–80. [[CrossRef](#)]
8. Demissie, T.A.; Sime, C.H. Assessment of the performance of CORDEX regional climate models in simulating rainfall and air temperature over southwest Ethiopia. *Heliyon* **2021**, *7*, e07791. [[CrossRef](#)]

9. Yuan, X.; Chen, C.; Lei, X.; Yuan, Y.; Muhammad Adnan, R. Monthly runoff forecasting based on LSTM–ALO model. *Stoch. Environ. Res. Risk Assess.* **2018**, *32*, 2199–2212. [\[CrossRef\]](#)
10. He, Y.; Chen, C.; Li, B.; Zhang, Z. Prediction of near-surface air temperature in glacier regions using ERA5 data and the random forest regression method. *Remote Sens. Appl. Soc. Environ.* **2022**, *28*, 100824.
11. Xiu, Y.; Wang, Q.; Li, Z.; Li, G.; Lu, P. Estimating spatial distributions of design air temperatures for ships and offshore structures in the Arctic Ocean. *Polar Sci.* **2022**, *34*, 100875. [\[CrossRef\]](#)
12. Ji, L.; Wang, Z.; Chen, M.; Fan, S.; Wang, Y.; Shen, Z. How much can AI techniques improve surface air temperature forecast?—A report from AI Challenger 2018 Global Weather Forecast Contest. *J. Meteorol. Res.* **2019**, *33*, 989–992. [\[CrossRef\]](#)
13. Cifuentes, J.; Marulanda, G.; Bello, A.; Reneses, J. Air temperature forecasting using machine learning techniques: A review. *Energies* **2020**, *13*, 4215. [\[CrossRef\]](#)
14. Mohammadi, B.; Mehdizadeh, S.; Ahmadi, F.; Lien, N.T.T.; Pham, Q.B. Developing hybrid time series and artificial intelligence models for estimating air temperatures. *Stoch. Environ. Res. Risk Assess.* **2021**, *35*, 1189–1204. [\[CrossRef\]](#)
15. Bayatvarkeshi, M.; Bhagat, S.K.; Mohammadi, K.; Kisi, O.; Farahani, M.; Hasani, A.; Deo, R.; Yaseen, Z.M. Modeling soil temperature using air temperature features in diverse climatic conditions with complementary machine learning models. *Comput. Electron. Agric.* **2021**, *185*, 106158. [\[CrossRef\]](#)
16. Ghadiri, M.; Marjani, A.; Mohammadinia, S.; Shirazian, S. An insight into the estimation of relative humidity of air using artificial intelligence schemes. *Environ. Dev. Sustain.* **2021**, *23*, 10194–10222. [\[CrossRef\]](#)
17. Meliho, M.; Khattabi, A.; Zejli, D.; Orlando, C.A.; Dansou, C.E. Artificial intelligence and remote sensing for spatial prediction of daily air temperature: Case study of Souss watershed of Morocco. *Geo-Spat. Inf. Sci.* **2022**, *25*, 244–258. [\[CrossRef\]](#)
18. Miyano, T.; Girosi, F. *Forecasting Global Temperature Variations by Neural Networks*; Technical Report; Massachusetts Institute of Technology, Cambridge Artificial Intelligence Laboratory: Cambridge, MA, USA, 1994.
19. Lanza, P.A.G.; Cosme, J.M.Z. A short-term temperature forecaster based on a state space neural network. *Eng. Appl. Artif. Intell.* **2002**, *15*, 459–464. [\[CrossRef\]](#)
20. Ustaoglu, B.; Cigizoglu, H.K.; Karaca, M. Forecast of daily mean, maximum and minimum temperature time series by three artificial neural network methods. *Meteorol. Appl. A J. Forecast. Pract. Appl. Train. Tech. Model.* **2008**, *15*, 431–445. [\[CrossRef\]](#)
21. Adnan, R.M.; Mostafa, R.R.; Islam, A.R.M.T.; Gorgij, A.D.; Kuriqi, A.; Kisi, O. Improving Drought Modeling Using Hybrid Random Vector Functional Link Methods. *Water* **2021**, *13*, 3379. [\[CrossRef\]](#)
22. Abhishek, K.; Singh, M.P.; Ghosh, S.; Anand, A. Weather forecasting model using artificial neural network. *Procedia Technol.* **2012**, *4*, 311–318. [\[CrossRef\]](#)
23. Jallal, M.A.; Chabaa, S.; El Yassini, A.; Zeroual, A.; Ibnyaich, S. Air temperature forecasting using artificial neural networks with delayed exogenous input. In Proceedings of the 2019 International Conference on Wireless Technologies, Embedded and Intelligent Systems (Wits), Fez, Morocco, 3–4 April 2019; pp. 1–6.
24. Tran, T.; Bateni, S.; Ki, S.; Vosoughifar, H. A review of neural networks for air temperature forecasting. *Water* **2021**, *13*, 1294. [\[CrossRef\]](#)
25. Paniagua-Tineo, A.; Salcedo-Sanz, S.; Casanova-Mateo, C.; Ortiz-García, E.G.; Cony, M.A.; Hernández-Martín, E. Prediction of daily maximum temperature using a support vector regression algorithm. *Renew. Energy* **2011**, *36*, 3054–3060. [\[CrossRef\]](#)
26. Ortiz-García, E.G.; Salcedo-Sanz, S.; Casanova-Mateo, C.; Paniagua-Tineo, A.; Portilla-Figueras, J.A. Accurate local very short-term temperature prediction based on synoptic situation Support Vector Regression banks. *Atmos. Res.* **2012**, *107*, 1–8. [\[CrossRef\]](#)
27. Mellit, A.; Pavan, A.M.; Benghamem, M. Least squares support vector machine for short-term prediction of meteorological time series. *Theor. Appl. Climatol.* **2013**, *111*, 297–307. [\[CrossRef\]](#)
28. Ikram, R.M.A.; Mostafa, R.R.; Chen, Z.; Islam, A.R.M.T.; Kisi, O.; Kuriqi, A.; Zounemat-Kermani, M. Advanced Hybrid Meta-heuristic Machine Learning Models Application for Reference Crop Evapotranspiration Prediction. *Agronomy* **2023**, *13*, 98. [\[CrossRef\]](#)
29. Chevalier, R.F.; Hoogenboom, G.; McClendon, R.W.; Paz, J.A. Support vector regression with reduced training sets for air temperature prediction: A comparison with artificial neural networks. *Neural Comput. Appl.* **2011**, *20*, 151–159. [\[CrossRef\]](#)
30. Salcedo-Sanz, S.; Deo, R.C.; Carro-Calvo, L.; Saavedra-Moreno, B. Monthly prediction of air temperature in Australia and New Zealand with machine learning algorithms. *Theor. Appl. Climatol.* **2016**, *125*, 13–25. [\[CrossRef\]](#)
31. Venkadesh, S.; Hoogenboom, G.; Potter, W.; McClendon, R. A genetic algorithm to refine input data selection for air temperature prediction using artificial neural networks. *Appl. Soft Comput.* **2013**, *13*, 2253–2260. [\[CrossRef\]](#)
32. Azad, A.; Kashi, H.; Farzin, S.; Singh, V.P.; Kisi, O.; Karami, H.; Sanikhani, H. Novel approaches for air temperature prediction: A comparison of four hybrid evolutionary fuzzy models. *Meteorol. Appl.* **2020**, *27*, e1817. [\[CrossRef\]](#)
33. Mallick, J.; Islam, A.R.M.T.; Ghose, B.; Islam, H.T.; Rana, Y.; Hu, Z.; Bhat, S.A.; Pal, S.C.; Ismail, Z.B. Spatiotemporal trends of temperature extremes in Bangladesh under changing climate using multi-statistical techniques. *Theor. Appl. Climatol.* **2022**, *147*, 307–324. [\[CrossRef\]](#)
34. Jahan, C.S.; Mazumder, Q.H.; Islam, A.T.M.M.; Adham, M.I. Impact of irrigation in Barind area, NW Bangladesh—An evaluation based on the meteorological parameters and fluctuation trend in groundwater table. *J. Geol. Soc. India* **2010**, *76*, 134–142. [\[CrossRef\]](#)

35. Kamruzzaman, M.; Rahman, A.T.M.S.; Kabir, M.E.; Jahan, C.S.; Mazumder, Q.H.; Rahman, M.S. Spatio-temporal analysis of climatic variables in the western part of Bangladesh. *Environ. Dev. Sustain.* **2016**, *18*, 89–108. [\[CrossRef\]](#)
36. Mostafa, R.R.; Kisi, O.; Adnan, R.M.; Sadeghifar, T.; Kuriqi, A. Modeling Potential Evapotranspiration by Improved Machine Learning Methods Using Limited Climatic Data. *Water* **2023**, *15*, 486. [\[CrossRef\]](#)
37. Pao, Y.H.; Takefuji, Y. Functional-link net computing: Theory, system architecture, and functionalities. *IEEE Comput.* **1992**, *25*, 76–79. [\[CrossRef\]](#)
38. Husmeier, D. Random vector functional link (RVFL) networks. In *Neural Networks for Conditional Probability Estimation*; Springer: London, UK, 1999; pp. 87–97.
39. Elsheikh, A.H.; Shehabeldeen, T.A.; Zhou, J.; Showaib, E.; Abd Elaziz, M. Prediction of laser cutting parameters for polymethyl-methacrylate sheets using random vector functional link network integrated with equilibrium optimizer. *J. Intell. Manuf.* **2021**, *32*, 1377–1388. [\[CrossRef\]](#)
40. Mirjalili, S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowl. -Based Syst.* **2015**, *89*, 228–249. [\[CrossRef\]](#)
41. Taher, M.A.; Kamel, S.; Jurado, F.; Ebeed, M. An improved moth-flame optimization algorithm for solving optimal power flow problem. *Int. Trans. Electr. Energy Syst.* **2019**, *29*, e2743. [\[CrossRef\]](#)
42. Li, Z.; Zhou, Y.; Zhang, S.; Song, J. Lévy-flight moth-flame algorithm for function optimization and engineering design problems. *Math. Probl. Eng.* **2016**, *2016*, 1423930. [\[CrossRef\]](#)
43. Savsani, V.; Tawhid, M.A. Non-dominated sorting moth flame optimization (NS-MFO) for multi-objective problems. *Eng. Appl. Artif. Intell.* **2017**, *63*, 20–32. [\[CrossRef\]](#)
44. Xu, L.; Li, Y.; Li, K.; Beng, G.H.; Jiang, Z.; Wang, C.; Liu, N. Enhanced moth-flame optimization based on cultural learning and Gaussian mutation. *J. Bionic Eng.* **2018**, *15*, 751–763. [\[CrossRef\]](#)
45. Khalilpourazari, S.; Khalilpourazary, S. An efficient hybrid algorithm based on Water Cycle and Moth-Flame Optimization algorithms for solving numerical and constrained engineering optimization problems. *Soft Comput.* **2019**, *23*, 1699–1722. [\[CrossRef\]](#)
46. Chen, C.; Wang, X.; Yu, H.; Wang, M.; Chen, H. Dealing with multi-modality using synthesis of Moth-flame optimizer with sine cosine mechanisms. *Math. Comput. Simul.* **2021**, *188*, 291–318. [\[CrossRef\]](#)
47. Kaur, K.; Singh, U.; Salgotra, R. An enhanced moth flame optimization. *Neural Comput. Appl.* **2020**, *32*, 2315–2349. [\[CrossRef\]](#)
48. Ma, L.; Wang, C.; Xie, N.G.; Shi, M.; Ye, Y.; Wang, L. Moth-flame optimization algorithm based on diversity and mutation strategy. *Appl. Intell.* **2021**, *51*, 5836–5872. [\[CrossRef\]](#)
49. Nadimi-Shahraki, M.H.; Fatahi, A.; Zamani, H.; Mirjalili, S.; Abualigah, L.; Abd Elaziz, M. Migration-based moth-flame optimization algorithm. *Processes* **2021**, *9*, 2276. [\[CrossRef\]](#)
50. Eskandar, H.; Sadollah, A.; Bahreininejad, A.; Hamdi, M. Water cycle algorithm—A novel metaheuristic optimization method for solving constrained engineering optimization problems. *Comput. Struct.* **2012**, *110*, 151–166. [\[CrossRef\]](#)
51. Messaoud, R.B.; Midouni, A.; Hajji, S. PEM fuel cell model parameters extraction based on moth-flame optimization. *Chem. Eng. Sci.* **2021**, *229*, 116100. [\[CrossRef\]](#)
52. 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. [\[CrossRef\]](#)
53. Alizamir, M.; Kisi, O.; Adnan, R.M.; Kuriqi, A. Modelling reference evapotranspiration by combining neuro-fuzzy and evolutionary strategies. *Acta Geophys.* **2020**, *68*, 1113–1126. [\[CrossRef\]](#)
54. Muhammad, R.; Zhongmin, A.; Kulwinder, L.; Parmar, S.; Soni, K.; Kisi, O.; Adnan, R.M. Modeling monthly streamflow in mountainous basin by MARS, GMDH- NN and DENFIS using hydroclimatic data. *Neural Comput. Appl.* **2020**, *33*, 2853–2871. [\[CrossRef\]](#)
55. Benaafi, M.; Yassin, M.A.; Usman, A.G.; Abba, S.I. Neurocomputing Modelling of Hydrochemical and Physical Properties of Groundwater Coupled with Spatial Clustering, GIS, and Statistical Techniques. *Sustainability* **2022**, *14*, 2250. [\[CrossRef\]](#)
56. Smith, B.A.; McClendon, R.W.; Hoogenboom, G. Improving Air Temperature Prediction with Artificial Neural Networks. *Int. J. Comput. Inf. Eng.* **2007**, *1*, 3159.
57. Meshram, S.G.; Pourghasemi, H.R.; Abba, S.I.; Alvandi, E.; Meshram, C.; Khedher, K.M. A comparative study between dynamic and soft computing models for sediment forecasting. *Soft Comput.* **2021**, *25*, 0123456789. [\[CrossRef\]](#)
58. Mirabbasi, R.; Kisi, O.; Sanikhani, H.; Gajbhiye, S. Monthly long-term rainfall estimation in Central India using M5Tree. *Neural Comput. Appl.* **2018**, *9*, 6843. [\[CrossRef\]](#)
59. Tikhamarine, Y.; Souag-Gamane, D.; Kisi, O. A new intelligent method for monthly streamflow prediction: Hybrid wavelet support vector regression based on grey wolf optimizer (WSVR-GWO). *Arab. J. Geosci.* **2019**, *12*, 540. [\[CrossRef\]](#)
60. Yaseen, Z.M.; Ebtehaj, I.; Kim, S.; Sanikhani, H.; Asadi, H.; Ghareb, M.I.; Bonakdari, H.; Mohtar, W.H.M.W.; Al-Ansari, N.; Shahid, S. Novel hybrid data-intelligence model for forecasting monthly rainfall with uncertainty analysis. *Water* **2019**, *11*, 502. [\[CrossRef\]](#)
61. Heddad, S.; Ptak, M.; Zhu, S. Modelling of daily lake surface water temperature from air temperature: Extremely randomized trees (ERT) versus Air2Water, MARS, M5Tree, RF and MLPNN. *J. Hydrol.* **2020**, *588*, 125130. [\[CrossRef\]](#)
62. Hinkel, K.M.; Paetold, F.; Nelson, F.E.; Bockheim, J.G. Patterns of soil temperature and moisture in the active layer and upper permafrost at Barrow, Alaska: 1993–1999. *Glob. Planet. Chang.* **2001**, *29*, 293–309. [\[CrossRef\]](#)

63. Singhal, M.; Gairola, A.C.; Singh, N. Artificial neural network-assisted glacier forefield soil temperature retrieval from temperature measurements. *Theor. Appl. Climatol.* **2021**, *143*, 1157–1166. [[CrossRef](#)]
64. Muhammad, R.; Andrea, A.; Kisi, O. Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stoch. Environ. Res. Risk Assess.* **2020**, *35*, 597–616. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.