Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2021.26397

215 (2021) 288–318
March

Do quadratic and Poisson regression models help to predict monthly rainfall?

Youssef Kassem[a,b,*], Hüseyin Gökçekuş[b]

[a]Department of Mechanical Engineering, Engineering Faculty, Near East University, 99138 Nicosia (via Mersin 10, Turkey), Cyprus, Tel. +90 (392) 2236464; emails: yousseuf.kassem@neu.edu.tr/youssef.kassem1986@hotmail.com (Y. Kassem)
[b]Department of Civil Engineering, Civil and Environmental Engineering Faculty, Near East University, 99138 Nicosia (via Mersin 10, Turkey), Cyprus, email: huseyin.gokcekus@neu.edu.tr (H. Gökçekuş)

Received 2 April 2020; Accepted 27 July 2020

ABSTRACT

Agricultural water scarcity in the primarily rainfed agricultural system of Jigawa State in Nigeria is more related to the variability of rainfall. Rainfed subsistence farming systems in this state generally obtain low crop yields and production as a result of highly erratic rainfall seasons. Thus, predicting rainfall in the region is of great significance as it could help the government to improve sustainable rainfed agriculture in the region. To enable the design of a model capable of accurate predictions, this paper summarizes recent scientific studies aimed at predicting rainfall in Nigeria and around the world utilizing artificial and mathematical models. According to this review, it is evident that quadratic and Poisson regression models have not yet been considered in other studies about monthly rainfall prediction. Additionally, few recent studies have used solar radiation and sunshine duration as input parameters for their models. Consequently, quadratic (QM) and Poisson regression (PRM) models are proposed for predicting the monthly rainfall in Jigawa State in the north–west of Nigeria. Monthly meteorological parameters including rainfall, average temperature, minimum temperature, maximum temperature, relative humidity, sunshine duration, solar radiation, and wind speed data spanning 10 y (2008–2017) obtained from the Nigerian Meteorological Agency were used in this study. Furthermore, temporal correlation and spatial correlation were applied to measure the relationship between monthly rainfall data and other meteorological parameters for the selected region. Moreover, the proposed models (QM and PRM) were compared with the most prominent rainfall artificial models (multilayer feed-forward neural network, cascade feed-forward neural network, and radial basis neural networks) to show the predictive accuracy of the proposed model. The results demonstrate that the developed PRM model is superior in predicting the value of monthly rainfall with reported values of 0.887 and 0.0542 for the parameters of R^2 and root mean squared error, respectively.

Keywords: Artificial models; Jigawa State; Monthly rainfall; Nigeria; Quadratic model; Poisson regression model

1. Introduction

Water use is affected by both changes in land use and farming intensity in already cultivated lands. Hence, to estimate the water needed to produce crop production, the conversion of rainfall into agricultural product effectiveness should be assessed. Several scientific studies have investigated the relationship between annual rainfall and production and losses that is essential for managing water in rain-fed farming [1,2]. Moreover, according to Dercon and Christiaensen [3], Falco and Chavas [4], and Amare et al. [5], rainfall is considered a direct input for the production of crops, and rainfall variability can affect agricultural productivity, that is, rainfall could lead to a change in crops, which could move production away from the planned production [6]. Therefore, sustainable agricultural production and

* Corresponding author.

Presented at the 2nd International Conference on the Environment Survival and Sustainability (ESS 2019), 7–11 October 2019, Nicosia, North Cyprus

climate change are interrelated processes [7]. Also, extreme weather events like warmer and drier conditions are associated with negative impacts on agricultural production [8].

Nigeria is likely to suffer increasing levels of climate change impacts because of its geographical location and weak institutional, human, economic, technological, and financial capacity to cope with the multiple impacts of these disruptions. Vulnerability to climate change is compounded by the over-dependence on climate-sensitive sectors, especially agriculture. In Nigeria, the agricultural sector contributed approximately 46% to the national gross domestic product in the third quarter of 2017 [9]. According to Olayide and Alabi [6] and Fjelde and Uexkull [10], the variability of rainfall affects the rain-fed agricultural and food production systems in the country, which led to affect the Nigerian economy.

1.1. Literature review related to predicting the rainfall

This work discusses the empirical models including artificial neural networks (ANNs) and mathematical models including the quadratic model; hence, it is important to briefly introduce the subject before commencing with the literature review. In recent years, empirical approaches like ANNs and multiple linear regressions (MLR) have been used as powerful modeling tools in the estimation of rainfall data. ANNs have emerged as a powerful technique for modeling complex functional relationships [11,12]. Moreover, the MLR is used to describe the relationship between two or more independent variables and one dependent variable [13]. Many studies have utilized ANNs and mathematical models to predict the hourly/daily/weakly/monthly rainfall in many countries around the world and mainly in Nigeria. Table 1 summarizes the key features of previous scientific studies. In general, a standard set of input parameters for predicting hourly/daily/weekly/monthly rainfall is not pre-established. The selection of parameters depends on the approaches used and the regions studied. The choice of a parameter may also be constrained by the availability of measured data. According to Table 1, it can be concluded that:

- Researchers have recently focused on modeling hourly/daily/weekly/monthly rainfall using artificial intelligence models.
- Researchers have utilized meteorological parameters such as minimum and maximum temperatures, wind speed, pressure, and relative humidity.
- Few studies have used climatological parameters like sunshine duration and solar radiation as input data for the empirical model to estimate the hourly/daily/weekly/monthly rainfall.

1.2. Scope of the present work

The findings of the literature review reveal a clear lack of monthly rainfall prediction models in Jigawa state in Nigeria. Furthermore, the previous studies related to Nigeria (Table 1) have shown interesting features such as the El Niño and La Niña phenomena on the meteorological conditions but few studies have used meteorological parameters such as minimum and maximum temperatures, wind speed, pressure, and relative humidity as input parameters for the empirical model. Finally, according to the authors' review, most previous works have used artificial intelligence models, ARIMA, and mathematical regression models in terms of least absolute deviation multiple regressions, multiple linear regression, cluster wise linear regression to estimate the hourly/daily/weekly/monthly rainfall. Moreover, only two references used solar radiation and sunshine duration as input parameters for empirical models. Therefore, the objectives of this paper are as follows:

- To investigate the link among monthly rainfall and meteorological parameters including average temperature, minimum temperature, maximum temperature, relative humidity, sunshine duration, solar radiation, and wind speed with regard to Jigawa state in Nigeria. In the current study, quadratic (QM) and Poisson regression (PRM) models are employed for the accurate prediction of monthly rainfall. The study uses monthly data for the period from 2008 to 2017, which was obtained from the Nigerian Meteorological Agency. The authors develop mathematical equations for predicting the monthly rainfall. These equations depend on average temperature, minimum temperature, maximum temperature, relative humidity, sunshine duration, solar radiation, and wind speed.
- To develop ANNs, namely multilayer feed-forward neural network (MFFNN), cascade feed-forward neural network (CFNN), and radial basis neural networks (RBNN), to predict the monthly rainfall of Jigawa state in Nigeria. The inputs of the model are monthly average temperature, minimum temperature, maximum temperature, sunshine duration, solar radiation, wind speed, and relative humidity. The optimum architecture for MFFNN and CFNN is constructed based on changing both the number of hidden neurons and topologies of the neural network during the training process. MATLAB is used to train, test, and validate the proposed artificial model.
- To compare the results obtained from the proposed models (QM and PRM) with those obtained using ANN models to show the superiority of the proposed models (QM and PRM).

2. Materials and methods

2.1. Case study

Jigawa State is located in the north–west geopolitical zone of Nigeria. It lies between latitudes 11.00°N to 13.00°N and longitudes 8.00°E to 10.15°E (see Fig. 1). The total landmass in the selected state is about 24,742 km^2. According to ground survey data, 14% of the total landmass is represented by wetlands (total wetlands size of 3,433.79 km). According to NPC, the population of Jigawa State is 5,041,500. The climate of the state is characterized by a long dry season and a short wet season. The average annual temperature is within the range of 21°C–31°C with an average of 25°C [41]. Also, total annual rainfall is varied from 600 to 1,000 mm [41]. The volume of surface water and groundwater in the

Table 1
Summary of researchers' studies in the modeling of rainfall by artificial neural networks and mathematical model

Reference	Location/country	Empirical approach	Input	Output
Purnomo et al. [14]	Ampel, Boyolali, Central Java	Artificial neural network	2001–2013 monthly rainfall rate data used as input	2014–2015 Monthly rainfall rate
Abdulkadir et al. [15]	Markurdi Ilorin, Lafia, Jos, Lokoja, Minna and Abuja, Nigeria	Artificial neural network	80% of monthly rainfall data used as input data	Monthly rainfall
Bagirov et al. [16]	Victoria, Australia	Cluster wise linear regression	Maximum temperature, minimum temperature, evaporation, vapor, and solar radiation	Monthly rainfall
Kashiwao et al. [17]	Japan	Multi-layer perceptron neural network (MLPNN) and radial basis function neural network	Atmospheric pressure, precipitation, temperature, vapor pressure, humidity, and wind speed	Total rainfall
Xiang et al. [18]	Kunming, Lincang, and Mengzi, Yunnan Province, China	Ensemble empirical mode decomposition – support vector machine – artificial neural network model	1951–2007 daily rainfall used as input data	2007–2015 daily rainfall used as output data
Mirabbasi et al. [19]	India	M5Tree model, multivariate adaptive regression spline, least-square support vector regression (LSSVR), a gene expressing programming (GEP), and artificial neural networks methods	Geographical information and periodicity	Rainfall
Zeynoddin et al. [20]	Pahang basin, Malaysia	Linear stochastic model – non-linear extreme learning machine method	70% of monthly rainfall used as input data	30% of monthly rainfall used as output data
Bello and Mamman [21]	Kano, Nigeria	Artificial neural network and Linear regression	Southern Oscillation Index; Nino1+2; Nino3; Nino3.4; Nino4	Monthly rainfall
Rodi et al. [22]	Peninsular Malaysia	Clonal selection algorithm	Temperature, relative humidity	Monthly rainfall
Hudnurkar and Rayavarapu [23]	India	Artificial neural network	Relative humidity, mean sea level pressure, maximum temperature, minimum temperature, average temperature, average wind speed, and wind direction	Daily summer monsoon rainfall

(Continued)

Table 1 Continued

Reference	Location/country	Empirical approach	Input	Output
Peter and Precious [24]	Bauchi, Nigeria	Multiple linear regression model and artificial neural network (ANN)	Monthly means of sea surface temperature, air temperature, specific humidity, relative humidity, and U-wind at a surface different pressure level	Seasonal rainfall
Chattopadhyay and Chattopadhyay [25]	India	Conjugate gradient descent learning-based back-propagation artificial neural network	1871–1971 average summer monsoon rainfall used as input data	1972–1999 average summer monsoon rainfall used as output data
Dash et al. [26]	India	*K*-nearest neighbor (KNN), artificial neural network (ANN), and extreme learning machine (ELM)	1871–2010 average summer monsoon rainfall used as input data	2011–2016 average summer monsoon rainfall used as output data
Mohammadpour et al. [27]	Iran	Artificial neural networks (ANNs), learning-cellular automation (LCA), and novel hybrid method of ANN and CLA	Temperature, humidity, wind speed, and pressure	Daily rainfall
Anh et al. [28]	Ca Mau province, Vietnam	Multilayer feed-forward neural network, Seasonal artificial neural network, ARIMA, and GA-SA models	85% of monthly rainfall data used as input data	Monthly rainfall
Ilaboya and Igbinedion [29]	Benin City, Nigeria	Multiple linear regression and artificial neural network	The monthly temperature, wind speed, relative humidity, vapour pressure	Seasonal rainfall
Velasco et al. [30]	Spain	Multilayer perceptron neural network	Average temperature, maximum temperature, minimum temperature, average wind speed, relative humidity, total rainfall, visibility, day, month	Week-ahead rainfall
Hossain et al. [31]	Western Australia	Multiple linear regression and artificial neural network	Lagged values of the oceanic climate drivers, El Niño Southern Oscillation and Indian Ocean Dipole	Seasonal rainfall
Lin et al. [32]	Taiwan	Hybrid gray model, autoregressive integrated moving average and artificial neural network	40 y annual maximum daily rainfall	10 years annual maximum daily rainfall

Ayodele and Precious [33]	Ikeja, Nigeria	Artificial neural network	Sea surface temperature (SST), U-wind at (surface, 700, 850, and 1,000), air temperature, specific humidity, ITD, and relative humidity	Seasonal rainfall
Reference	Location/country	Empirical approach	Input	Output
Bensafi et al. [34]	Setif, Algeria	*K* nearest weighted neighbors (WkNN) classification algorithm	70% of the data used as input	Instantaneous rainfall
Pour et al. [35]	Peninsular Malaysia	Support vector machines (SVM), random forests (RF), and Bayesian artificial neural networks	Rainfall amount, average rainfall intensity, days with rainfall more than 95th percentile rainfall, and dry days	Seasonal rainfall and rainfall extremes
Pham et al. [36]	Vietnam	Adaptive network-based fuzzy inference system optimized with particle swarm optimization, artificial neural networks, and support vector machines	Maximum temperature, minimum temperature, wind speed, relative humidity, and solar radiation	Daily rainfall
Ali et al. [37]	Pakistan	Complete ensemble empirical mode decomposition (CEEMD) combined with Random Forest (RF) and Kernel ridge regression (KRR) algorithms in designing a hybrid CEEMD-RF-KRR model	Rainfall data were collected from three stations	Monthly rainfall
Gökçekuş et al. [38]	Morphou, Northern Cyprus	Artificial neural networks	Minimum temperature, maximum temperature, average temperature, wind speed, global solar radiation, and sunshine duration	Monthly rainfall
Diop et al. [39]	Senegal	Multilayer perceptron-whale optimization algorithm (MLP-WOA)	75% of the data used as input	Annual rainfall
Chong et al. [40]	Peninsular Malaysia	Convolutional neural network and wavelet transform	80% of the data used as input	Monthly rainfall and daily rainfall

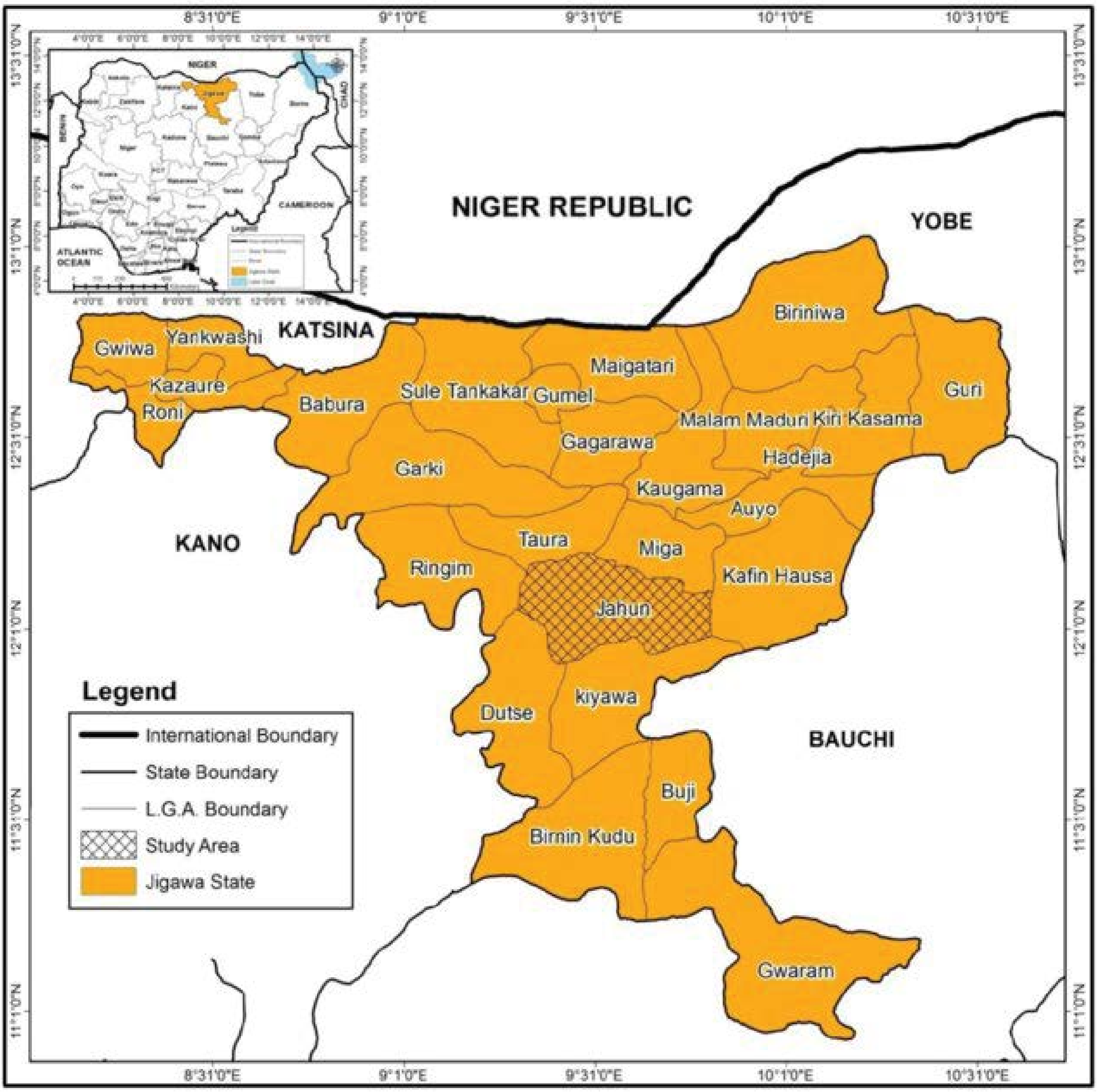


Fig. 1. Jigawa State map.

state is about 477 mcm and 30,000–40,000 m^3/km^2, respectively and the annual water recharge from rainfall is around 3,676 mcm. Agriculture is major economic activity in the state. Hence, the state has a high potential for both production and consumption compared to other states in Nigeria. Jigawa as a state has a large expanse of agricultural land, rivers, and flood plains suitable for crops, livestock, and fish production. However, Jigawa state faces significant water challenges. The effect of climate change on water resources in the state should be addressed in terms of its relation to the water cycle, water pollution, water scarcity, poor water administration, lack of resources for research and technological development, and lack of environmental planning. The agriculture in Jigawa state is divided into two farming systems, namely rainfed agriculture (agriculture that relies on natural rains) and irrigated farming systems. The total land area of the state is 2.24 million ha, out of which 1.6 million ha are estimated to be cultivated during the raining season, while over 400,000 ha of the landmass has potential for irrigated cultivation. Therefore, the rainfed agriculture system is considered as the primary system in the state, which occupies more than 75% of the cultivated land. It is considered the most important and traditional cultivation method that is dependent on rainwater. Jigawa state depends on a rainfed agriculture system for the widespread production of millet, sorghum, cowpea, groundnuts, sesame, rice, and maize, while irrigated farming is used for the production of tomatoes, pepper, onions, wheat, sugarcane, okra, carrot, lettuce, maize, and a host of other leafy vegetables. Moreover, agriculture provides a livelihood for

90% of the population and is the main employer in the state. Thus, rural livelihoods are strongly dependent on rainfall patterns and the frequent dry spells during cropping season's impact negatively on food security. This does not mean that rainfed agriculture is a problem in itself, but that it is more vulnerable to risks and tends to be less productive.

2.2. Data collected

To investigate the link among rainfall and meteorological parameters including average temperature, minimum and maximum temperatures, relative humidity, solar radiation, wind speed, and sunshine duration as related to the selected region, the study uses monthly data for the period from 2008 to 2017, which are collected from the Nigerian Meteorological Agency. The data are measured at various heights.

2.3. Statistical analysis

Rainfall analysis depends on its distribution pattern. In the literature, different distribution functions such as Gumbel, Weibull, Gamma, and Pearson type distribution are utilized to analyze the rainfall characteristics in specific regions [42,43]. In this study, five different distribution functions are chosen to select the best distribution function for analyzing the rainfall characteristics of the selected region. Skewness and Kurtosis values are used to find the best fitting probability distribution function to monthly rainfall, as shown in Table 2.

2.4. Correlation analysis

In the current study, Pearson product-moment [45] correlation is used to investigate the relationship between the rainfall (R) and meteorological parameters including mean temperature (T_{avg}), maximum temperature (T_{max}), minimum temperature (T_{min}), wind speed (WS), solar radiation (SR), sunshine duration (SD), and relative humidity (RH) followed by a parametric method for normal distribution. Additionally, in spatial correlation, meteorological parameters are utilized for correlation with the rainfall data followed by a non-parametric method for non-normal distribution series (Spearman correlation coefficient) to test the spatial statistical significance of the results [45]. Correlation coefficients and P-values are calculated using Minitab 17 software.

2.4. Empirical models

Generally, the method of selecting the models depends on the goal of the study. In this study, three artificial intelligence models including cascade forward neural network (CFNN), MFFNN, and radial basis function neural networks (RBFNN) are employed to estimate the monthly rainfall, which are compared with the quadratic model (QM).

2.4.1. Artificial intelligence models

The ANN is a powerful mathematical modeling tool especially for complex systems [46]. ANNs have long been used as an alternative methodology in different areas such as function approximation and so on. Many types of ANNs have been developed by scientific researchers such as of which the MFFNN is one of the most popular ANNs. The node numbers in the input and output layers are estimated by the nature of the problem. In this study, MFFNN, RBFNN, and CFNN are used for predicting the monthly rainfall in the selected study.

2.4.1.1. Multilayer feed-forward neural network

In general, MFFNN consists of input layer, one or two hidden layers, and output layer. In the present study, the input layer is composed of mean temperature, minimum and maximum temperatures, relative humidity, solar radiation, wind speed, and sunshine duration. The output layer has one node, which is the monthly rainfall. In order to determine the optimum number of the node in the hidden layer, a trial, and error approach is used. In this work, TRAINLM is used as a training function that updates the weight and bias values of the neuron connections according to Levenberg–Marquardt (LM) optimization. The back-propagation algorithm is used as a learning algorithm and it is a gradient descent algorithm. The logistic-sigmoid (logsig) and tangent-sigmoid (tansig) are used as activation functions whose outputs lie between 0 and 1 and are defined as:

$$\text{logsig} = \frac{1}{1+e^{-x}} \quad (1)$$

$$\text{tansig} = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \quad (2)$$

Table 2
Distribution curve selections [44]

Distribution type number	Distribution curve	Skewness (S) range	Kurtosis (K) range
I	Normal	$-0.4 < S < 0.4$	$-0.8 < K < 0.8$
II	Almost normal with positive tail	$S \geq 0.4$	$-0.8 < K < 0.8$
III	Narrow peak with positive tail	$S \geq 0.4$	$K \leq -0.8$ $K \geq 0.8$
IV	Almost normal with negative tail	$S \leq -0.4$	$-0.8 < K < 0.8$
V	Narrow peak with negative tail	$S \leq -0.4$	$K \geq 0.8$
VI	Bimodal, symmetrical with flat peak	$-0.4 < S < 0.4$	$K \leq -0.8$

The key step for developing an ANN is the training procedure, where the weights and biases are adjusted to minimize the difference between the output of the ANN and the actual value. In order to find the best performance for the ANN trained model, the mean squared error (MSE) is used. Fig. 2 presents the prediction processes used the proposed MFFNN method.

2.4.1.2. Cascade feed-forward neural network

Cascade-forward networks are similar to feed-forward networks (FFNN), but include a connection from the input and every previous layer to the following layers [47–49]. In this paper, the inputs are mean temperature, minimum and maximum temperatures, relative humidity, solar radiation, wind speed, and sunshine duration. The output is the monthly rainfall. Trial and error method is used to calculate the number of neurons. In this case, the number of hidden neurons should be lower than the optimal numbers, that is, if the number of hidden neurons is higher than the optimal number, then over-fitting and high variance may occur. Thus, using the iterative method to determine the optimum number of neurons based on the minimum value of root mean squared error (RMSE). The flowchart describing the steps of the proposed CFNN based method for predicting monthly rainfall is given by Fig. 3.

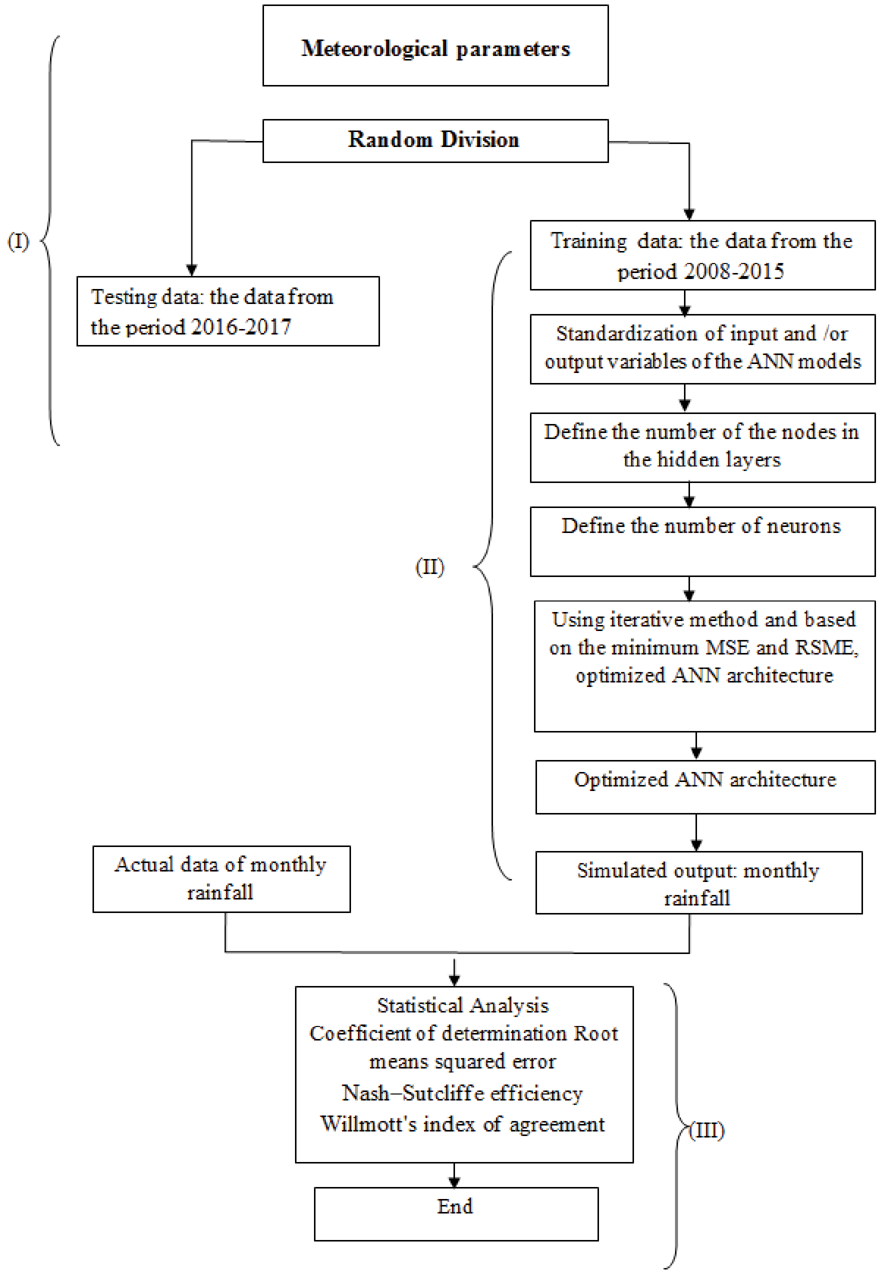


Fig. 2. Flowchart of the MFFNN based method prediction procedure.

2.4.1.3. Radial basis neural networks

RBFNN is one of the most popular kinds of ANNs that utilizes radial basis functions as activation functions [50]. It is a type of FFNN composed of three layers (input, hidden, and output layers). In this paper, Gaussian function is used as the transfer function in computational units. Also, the training of the RBFNN model is terminated once the calculated error reached the desired values or number of training iterations. The number of nodes of input layer is identical to the number of model inputs. The steps are illustrated in Fig. 4 to provide a better performance for implemented RBFNN model. In this work, the data from the period 2008–2015 are used for training and the rest of the data (2016–2017) are utilized to test.

2.4.2. Quadratic model

A QM is a mathematical model that represents a simple description of a physical, chemical, or biological process. The quadratic model of monthly rainfall is developed using a response surface methodology (RSM). The mathematical-quadratic-model is used to investigate the influence of interactive effects of the meteorological parameters (average temperature (T_{avg}), minimum (T_{min}) and maximum (T_{max}) temperatures, relative humidity (RH), solar radiation (SR), wind speed (WS), and sunshine duration (SD) on monthly rainfall (R). In the RSM method, the quantitative form of the relationship between the independent input variables and desired output is expressed as follows:

$$R = f\left(T_{avg}, T_{min}, T_{max}, \mathrm{SR}, \mathrm{SD}, \mathrm{WS}, \mathrm{RH}\right) \quad (3)$$

On the basis of the actual data, regression analysis was carried out by the following quadratic polynomial model:

$$R = \beta_0 + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \beta_{ii} x_i^2 + \sum_{i}^{n-1} \sum_{i=i+1}^{n} \beta_{ij} x_i x_j \quad (4)$$

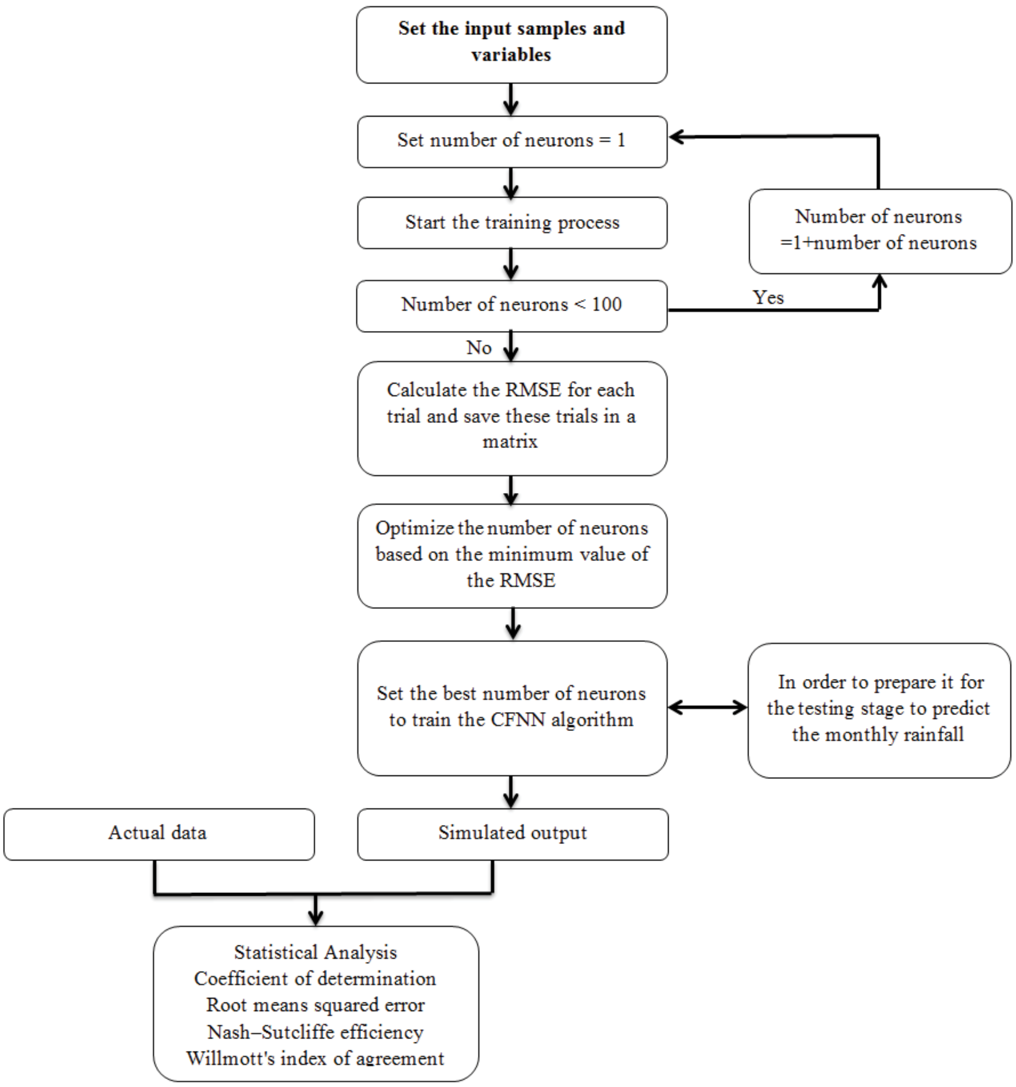


Fig. 3. Proposed algorithm of predicting monthly rainfall using CFNN.

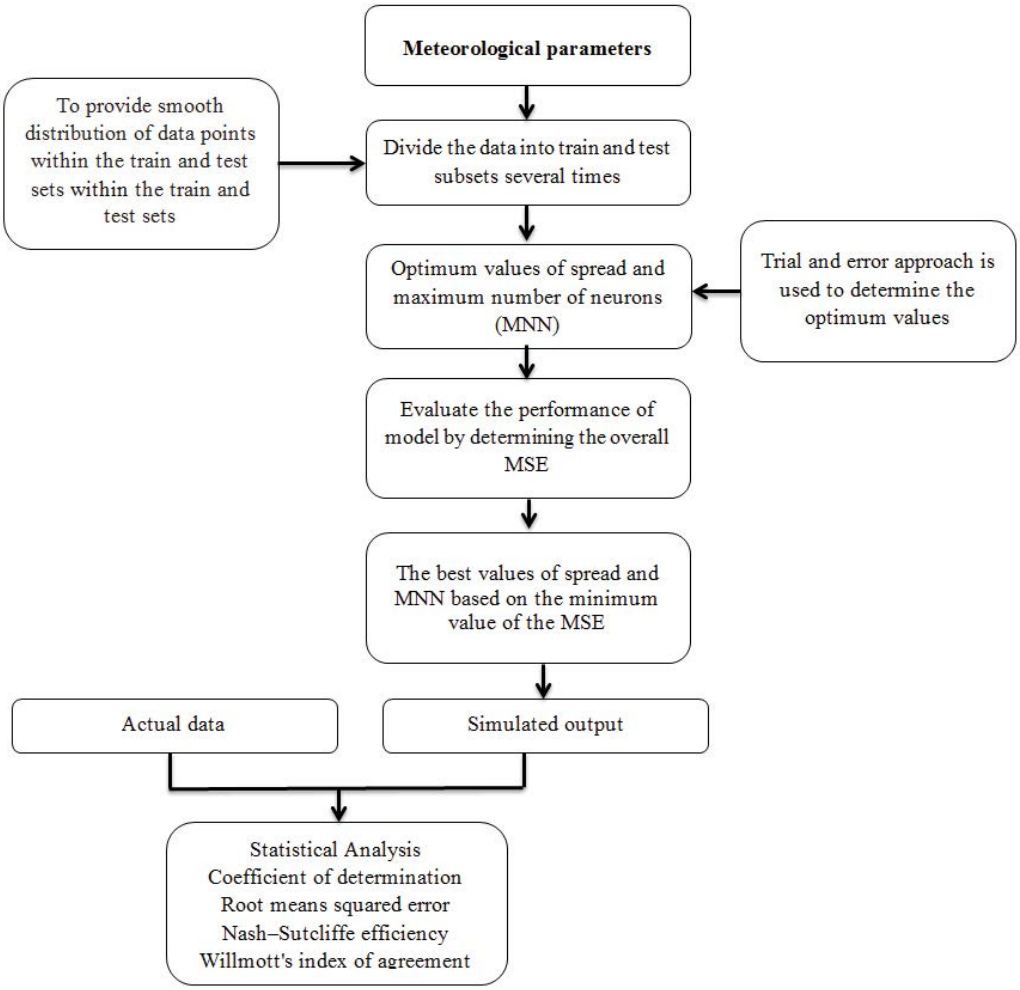


Fig. 4. Proposed algorithm of predicting monthly rainfall using RBFNN.

where β_0 is the offset term; β_i is the linear coefficient; the second-order coefficient and β_{ij} is the interaction coefficient; x_i and x_j are the independent variables. The least squares method was employed to ascertain the values of the model parameters and analysis of variance (ANOVA) was applied to establish their statistical significance at a confidence level of 95%.

2.4.3. Poisson regression model

Poisson regression is a generalized linear model (GLM) commonly used to model rare events and count data. A large number of academicians in many different fields have used PRM in their studies [51]. There are two main assumptions made when using Poisson regression. The first is that the response variable follows a Poisson distribution.

$$P = e^{\lambda} \frac{\lambda^k}{k!} \tag{5}$$

where P is the probability that k number of events will occur per interval of time and λ is the event rate. The second major assumption when using Poisson regression is that the variance and the mean of the response variable are equal. Thus, the probability distribution (Eq. (8)) can be specified by only one parameter, λ [52].

In Poisson regression, the mean parameter, λ is defined by the log-linear function:

$$\lambda = \exp\left(-x_i \beta\right) \tag{6}$$

where x_i is a vector of input values for time i and β is a corresponding vector of model parameters, which is optimized during training [52].

2.5. Model performance criteria

In general, the performance measures are utilized to select the "better" predictive model. The following statistical indicators are widely used to assess the predictive power of ANN and mathematical models [53,54].

Coefficient of determination (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2}{\sum_{i=1}^{n}\left(a_{p,i} - a_{a,\text{ave}}\right)^2} \tag{7}$$

MSE:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2 \quad (8)$$

RMSE:

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2} \quad (9)$$

Nash–Sutcliffe efficiency (NSE):

$$\text{NSE} = 1 - \frac{\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2}{\sum_{i=1}^{n}\left(a_{a,i} - a_{a,\text{ave}}\right)^2} \quad (10)$$

Willmott's index of agreement (d):

$$d = 1 - \frac{\sum_{i=1}^{n}\left(a_{a,i} - a_{p,i}\right)^2}{\sum_{i=1}^{n}\left(\left|a_{p,i} - a_{a,\text{ave}}\right| + \left|a_{a,i} - a_{a,\text{ave}}\right|\right)^2} \quad (11)$$

where n is the number of data, $a_{p,i}$ is the predicted values, $a_{a,i}$ is the actual values, $a_{a,\text{ave}}$ is the average actual values, and i is the number of input variables.

3. Results and discussions

3.1. Analysis of measurement data

In this section, the monthly rainfall (R) data are analyzed statistically. The statistical characteristics including arithmetic mean (mean) standard deviation (SD), coefficient of variation in percent (CV), minimum (min), the first and third quartiles (Q1 and Q3), median, maximum (max), skewness (S), and kurtosis (K) of the monthly rainfall for the selected region are summarized in Table 3. In addition, the table shows the type of distribution (DT) for each year. It is found that the mean values of monthly rainfall are within the range of 66.7–147.1 mm. The maximum value of monthly rainfall occurred in August 2012 with a value of 646.9 mm (Fig. 3) and the minimum value of 257 mm (Fig. 5) was recorded in August 2017. According to Kassem and Gökçekuş [43], to select the best distribution function, the skewness and kurtosis values are used to select the type of distribution. Therefore, a narrow peak with positive tail frequency distribution curves characterizes the mean monthly rainfall for all years (Table 2). Based on Fig. 2, it is observed that the highest average amount of rainfall is recorded in August with a value of 396.03 mm followed by July with a value of 288.94 mm.

Moreover, the monthly variations in the meteorological parameters including mean temperature (T_{avg}), maximum temperature (T_{max}), minimum temperature (T_{min}), wind speed (WS), solar radiation (SR), sunshine duration (SD), and relative humidity (RH) are shown in Fig. 6. It is found that the mean temperature values range between 17.54°C and 33.57°C. Also, it is noticed that the monthly maximum temperature was recorded in May 2015 with a value of 41.07°C, while the monthly minimum temperature of 8.37°C was obtained in January 2014. Moreover, as shown in Fig. 6, the monthly variations in the solar radiation and sunshine duration are within the range of 15.8–22.6 MJ/m^2/d and 4.7–10.4 h/d, respectively. Also, it is found that the selected region has high wind speeds, which vary from 2.2 to 7.5 m/s. Table 4 lists the annual meteorological parameters for the selected region during the investigation period of 2008–2017.

3.2. Correlation analysis

According to the data characteristics, the temporal correlation, and spatial correlation coefficient between the rainfall and meteorological parameters are computed and tabulated in Tables 5 and 6. According to Table S1, there are significant correlation coefficients between the minimum and maximum temperatures and rainfall for all periods. Also, it is observed that there is a significant positive relationship between the monthly temperatures and rainfall. According to Table 6, there are significant correlation coefficients between relative humidity and rainfall for all periods. Furthermore, the spatial correlation analysis between the temperatures in terms of average temperature, minimum temperature and maximum temperature, and rainfall

Table 3
Statistical estimators of the mean monthly rainfall for the period 2008–2017

Year	Mean	SD	CV	Minimum	Q1	Median	Q3	Maximum	S	K	DT
2008	86.3	140.2	162.4	0	0	4	165.4	421.7	1.61	1.86	III
2009	82.7	136.1	164.61	0	0	1.1	158.2	376	1.55	1.05	III
2010	91.5	114.1	124.74	0	0	50.5	219.6	291.5	0.99	–0.73	III
2011	101	130.9	129.6	0	0	23.4	217	378.9	1.06	–0.02	III
2012	147.1	231.2	157.18	0	0	9.8	358.2	646.9	1.41	0.54	III
2013	76.5	130.7	170.78	0	0	12.6	137.4	440.1	2.29	5.6	III
2014	116.3	183.8	157.96	0	0	23.9	166.7	509.6	1.68	1.57	III
2015	75.7	148.4	196.02	0	0	0	119	514.3	2.72	8.08	III
2016	79.2	112.6	142.05	0	0	5.5	187.9	323.9	1.24	0.37	III
2017	66.7	93.8	140.58	0	0	1.3	159	257	1.04	–0.43	III

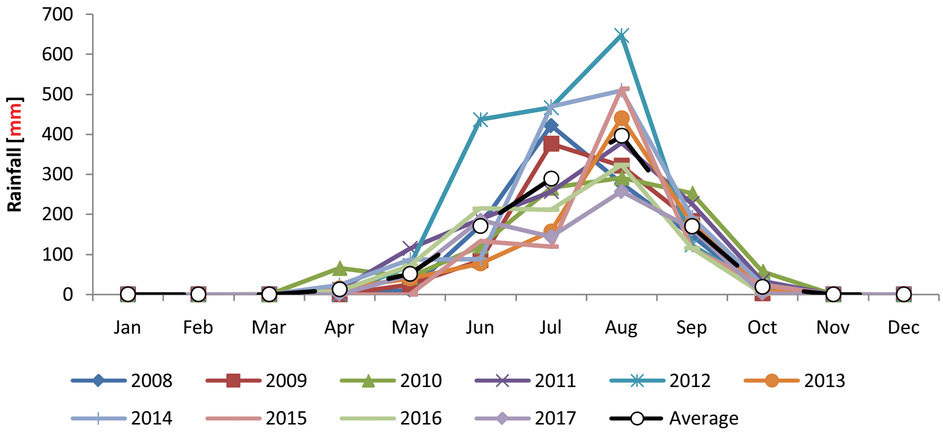


Fig. 5. Monthly rainfall during the investigation period 2008–2017.

shows a positive effect between the temperatures and rainfall, as shown in Table S2. The same results have been found by Gökçekuş et al. [38]. The authors concluded that temperature is considered as the most important parameter that has a greater impact on the estimated rainfall. It is found that wind speed has a minimum effect on rainfall prediction.

3.3. Artificial models

As mentioned previously, three neural network models were employed to predict the monthly rainfall for the selected region. Various independent variables are considered as inputs, as shown in Table 5. From the given data (2008–2017), the data from the period 2008–2015 are used for training and the rest of the data (2016–2017) are utilized to test the model. A series of models are examined to estimate the optimum number of hidden layers (HL), number of neurons (NN), and transfer function (TF) for the MFFNN and CFNN models, as shown in Table 6 for some of the trial and error iterations performed for evaluating the best HL, NN, and TF. It should be noted that the number of HLs and NNs in the MFFNN, CFNN models were determined by utilizing trial and error approaches. Based on the value of MSE, it is found that model MFFNN-VI and model CFNN-II have the minimum MSE value compared to other models. Hence, model MFFNN-VI and model CFNN-II are chosen as the best training model to predict the monthly rainfall due to the values of MSE and RMSE. Furthermore, the results show that MFFNN and CFNN with logsig was the most successful learning algorithm for the estimation of monthly rainfall.

Moreover, the 10-th order root of the input data was used instead of actual input data in order to provide better performance for the RBFNN model. This helps to smooth the variation of the input data points within a narrower range and this leads to better accuracy of the implemented model. Then, the data points were randomly divided into training and testing subsets. The random division was carried out several times to prevent aggregation of data points in the desired domain of the problem and to provide a smooth distribution of data points within the training and testing sets. In general, the spread and the maximum number of neurons (MNN) are important parameters in the structure of RBFNN as the performance and accuracy of the implemented model are significantly affected by the values of these parameters. Similarly, the optimum values of these parameters were estimated by the trial and error approach. Table 6 lists some of the trial and error iterations performed for evaluating the best values of spread and MNN. As is indicated in Table 6, the optimum values that provide the most accurate performance for the RBFNN model are 0.001 and 200 for the spread and MNN, respectively (RBFNN-VIII).

The comparison between the actual data used for the training and the data computed by the best ANN models is shown in Fig. S1. In addition, R^2 is used to evaluate the performance of artificial models. R^2 means the degree of the linear relationship between the observed and modeled values. The line is almost straight with a 45° angle and this proves the accuracy of the provided model. For the training phase, the R^2 values were found to be 0.910, 0.9195, and 0.9087 for MFFNN, CFNN, and RBFNN, respectively, as shown in Fig. S1. The results obtained from the ANN models show that the use of ANN is enough to predict monthly rainfall.

Furthermore, the actual data used for the testing and simulated data obtained from the best ANN models are compared through a linear regression model, as shown in Figs. S2–S4. The results indicate that MFFNN-VI has a higher R^2 compared to the other artificial models, as shown in Fig. 6. On the other hand, it should be noted that a higher R^2 value does not guarantee that the former is better than the latter because R^2 is a measure of the degree of the linear relationship between the actual and estimated values. Therefore, to select the best model for predicting monthly rainfall, RMSE is calculated. It is found that MFFNN-VI has a lower RMSE compared to the other models as shown in Table 6.

3.4. Mathematical models

The developed mathematical models (MM) including PRM and QM were implemented to predict the monthly value of rainfall for the selected region in Nigeria during the investigation period of 2008–2017. Fig. 7 illustrates the proposed procedure used in the PRM and QM models for predicting the monthly rainfall. The data of T_{avg}, T_{min}, T_{max},

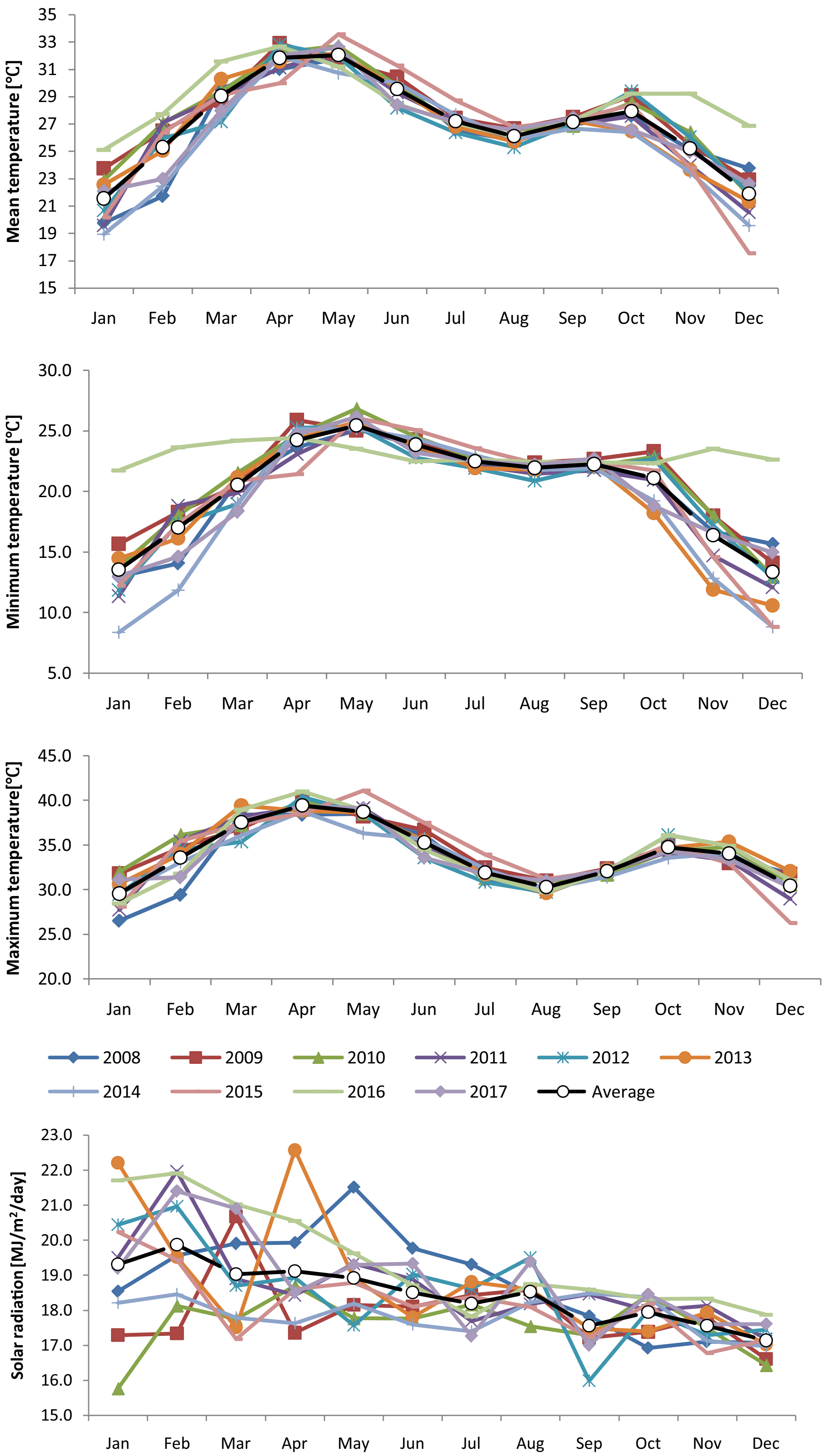


Fig. 6. Continued

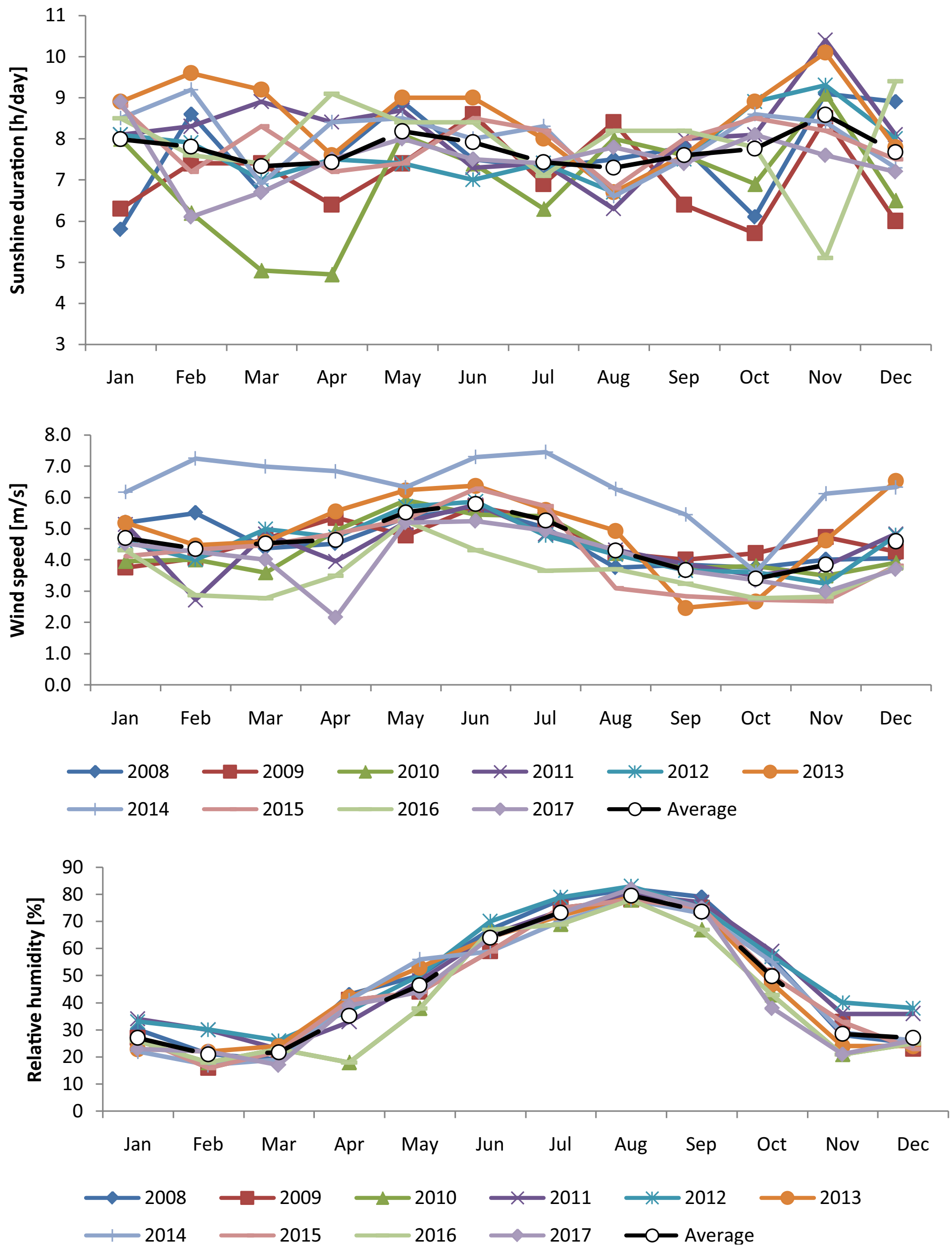


Fig. 6. Monthly variation of meteorological parameters.

SD, SR, WS, RH, and M were used to generate a mathematical equation based on PRM and QM for R as given in Eqs. (15) and (16), respectively.

$$R=\exp\begin{pmatrix}-166.1+0.001031\cdot M+6.241\cdot T_{avg}+1.0362\cdot \mathrm{RH}+\\ 7.468\cdot \mathrm{SD}+0.166\cdot \mathrm{WS}+2.191\cdot \mathrm{SR}-0.177\cdot T_{min}-\\ 0.0344\cdot T_{avg}^2-0.001805\cdot \mathrm{RH}^2+0.1603\cdot \mathrm{SD}^2-\\ 0.1172\cdot \mathrm{SR}^2-0.0435\cdot T_{min}^2-0.03521\cdot T_{max}^2-\\ 0.02984\cdot T_{avg}\cdot RH-0.0916\cdot T_{avg}\cdot SD+0.1562\cdot\\ T_{avg}\cdot \mathrm{SR}-0.01182 RH\cdot \mathrm{SD}+0.00964\cdot \mathrm{RH}\cdot\\ \mathrm{SR}-0.3624\cdot D\cdot \mathrm{SR}\end{pmatrix} \tag{12}$$

$$\begin{aligned}R={}&79.9+31.6\cdot M-146\cdot T_{avg}+166.5\cdot \mathrm{RH}-98.7\cdot \mathrm{SD}+\\ &38.1\cdot \mathrm{WS}+57.7\cdot \mathrm{SR}-130\cdot T_{min}-50.6\cdot M^2-21\cdot T_{avg}^2+\\ &109\cdot \mathrm{RH}^2+19.3\cdot \mathrm{SD}^2+32.2\cdot \mathrm{WS}^2-32.9\cdot \mathrm{SR}^2+\\ &7\cdot T_{min}^2+35\cdot T_{max}^2+123\cdot M\cdot T_{avg}+37.6\cdot M\cdot \mathrm{RH}-\\ &14.5\cdot M\cdot \mathrm{SD}-7.7\cdot M\cdot \mathrm{WS}+10.3\cdot M\cdot \mathrm{SR}-\\ &156\cdot M\cdot T_{min}-149\cdot T_{avg}\cdot \mathrm{RH}-96\cdot T_{avg}\cdot \mathrm{SD}-\\ &11\cdot T_{avg}\cdot \mathrm{WS}-49\cdot T_{avg}\cdot \mathrm{SR}-106.5\cdot \mathrm{RH}\cdot \mathrm{SD}+\\ &76.3\cdot \mathrm{RH}\cdot \mathrm{WS}+86.8\cdot \mathrm{RH}\cdot \mathrm{SR}-114\cdot \mathrm{RH}\cdot T_{min}+\\ &1.5\cdot \mathrm{SD}\cdot \mathrm{WS}-58.7\cdot \mathrm{SD}\cdot \mathrm{SR}+112\cdot \mathrm{SD}\cdot T_{min}-\\ &4.9\cdot \mathrm{WS}\cdot \mathrm{SR}+37\cdot \mathrm{WS}\cdot T_{min}+67\cdot \mathrm{SR}\cdot T_{min}\end{aligned} \tag{13}$$

Table 4
Annual meteorological parameters for the selected region

Year	T_{avg} (°C)	T_{min} (°C)	T_{max} (°C)	RH (%)	WS (m/s)	SD (h/d)	SR (MJ/m²/d)	R (mm)
2008	26.68	20.04	33.31	48.17	4.59	7.64	18.83	1,035.9
2009	27.75	21.04	34.46	45.25	4.59	7.12	17.91	992.2
2010	27.61	20.84	34.37	41.08	4.37	6.97	17.60	1,097.6
2011	26.70	19.60	33.79	49.58	4.40	8.17	18.72	1,211.6
2012	26.90	19.92	33.87	51.50	4.51	7.73	18.54	1,764.7
2013	26.87	19.35	34.39	45.67	4.94	8.53	18.81	918.2
2014	25.94	18.44	33.43	45.42	6.34	8.02	17.87	1,396
2015	26.91	19.70	34.12	45.25	4.18	7.88	18.18	908.3
2016	28.53	22.99	34.08	41.08	3.58	7.93	19.43	950.9
2017	26.78	19.80	33.76	43.67	4.03	7.52	18.83	800.4

Table 5
Study variables and their explanations

Parameters	Variable	Explanation	Limit		Unit
			Minimum	Maximum	
Input 1	M	Month	1	120	–
Input 2	T_{avg}	Average temperature	32.90	18.93	°C
Input 3	T_{min}	Minimum temperature	26.82	8.37	°C
Input 4	T_{max}	Maximum temperature	40.96	26.52	°C
Input 5	WS	Wind speed	7.46	2.16	m/s
Input 6	SD	Sunshine duration	10.4	4.7	h/d
Input 7	SR	Solar radiation	22.57	15.76	MJ/m²/d
Input 8	RH	Relative humidity	83	16	%
Output 1	R	Rainfall	646.9	0	mm

The results of the actual data and the corresponding values predicted by Eqs. (15) and (16) are displayed in Figs. S5 and S6. To test the fit of the model, R^2 is determined. For higher modeling accuracy, the R^2 value should be closer to 1. In this case, the values of R^2 for testing data are 0.8867 for PRM and 0.8786 for QM.

3.5. Performance evaluation of artificial models and mathematical models for testing data

As mentioned previously, to compare the performance of the models, data for 2008–2015 are used as the training part and those from 2016 to 2017 are used to test each model. Furthermore, the R^2, RMSE, NSE, and Willmott's index of agreement (d) are determined in order to select the best model for predicting monthly rainfall. R^2 is a measure of how well the regression line represents the data, while RMSE is a direct method for describing deviations. For high accuracy, R^2 must be close to 1.0 and the RMSE between the observed and predicted values must be as small as possible. Table 7 shows the results of the R^2 and RMSE values for all models. It is observed that all models gave good predictions according to the R^2 and RMSE values for the testing data. Also, it is found that the mathematical models including PRM and QM have the highest value of R^2 and lowest value of RMSE for the testing data. By comparing the computation results, the fitting precision of PRM model is higher than those of other models, where the highest R^2 and least RMSE are 0.887 and 0.0542, respectively.

Moreover, the NSE is generally similar to the R^2 measure for goodness-of-fit. A value of NSC = 1 indicates perfectly good forecasting accuracy; NSE = 0 when a forecast is no better than using the mean of the observed data; and NSE has negative values when a forecast is less accurate than the reference forecast. Thus, it is found that the NSE values for the ANN models show that the MFFNN, RBFNN models are satisfactory but the CFNN model is not satisfactory. Additionally, the NSE values for the PRM and QM models are 0.875 and 0.850, respectively, which indicate that they are acceptable as shown in Table 7.

Furthermore, the performance of the predictive models is evaluated using Willmott's index of agreement (d). Willmott's index of agreement (WIA) is standard measure to determine the error degree of the model. As shown in Table 7, it is found that the PRM model is the best and that the other models of monthly rainfall are also acceptable. Moreover, Fig. 8 shows the statistical indicators (R^2, RMSE, NSE, and d) for better comparison between the models.

4. Discussions

The findings of this study are important for agricultural production and other socio-economic activities, which are directly concerned with the rainfed agricultural system.

Table 6
Evaluation of the networks and statistical tool's performance of the artificial models

ANN model	Model number	NH	NN	TF	MSE-training	RMSE-Training	RMSE-Testing
MFFNN	MFFNN-I	1	5	logsig	0.00584	0.07639	0.07842
	MFFNN-II	1	8	logsig	0.00969	0.09843	0.13454
	MFFNN-III	1	10	logsig	0.00569	0.05539	0.09002
	MFFNN-IV	1	5	tansig	0.00555	0.07451	0.11902
	MFFNN-V	1	8	tansig	0.01189	0.10905	0.23940
	MFFNN-VI	1	12	logsig	0.00483	0.06951	0.06030
	MFFNN-VII	2	5	logsig	0.01402	0.11839	0.09793
	Model number	NH	NN	TF	MSE-training	RMSE-Training	RMSE-Testing
CFNN	CFNN-I	1	5	logsig	0.00479	0.06922	0.14237
	CFNN-II	1	8	logsig	0.00442	0.06645	0.16377
	CFNN-III	1	10	logsig	0.00846	0.09200	0.08677
	CFNN-IV	1	5	tansig	0.00577	0.07595	0.12205
	CFNN-V	2	5	logsig	0.00605	0.07780	0.09080
	Model number	Spread	MNN	–	MSE-training	RMSE-Training	RMSE-Testing
RBFNN	RBFNN-I	0.01	100	–	0.00902	0.09496	0.14643
	RBFNN-II	0.01	10	–	0.00997	0.09983	0.09745
	RBFNN-III	0.01	50	–	0.00902	0.09496	0.14643
	RBFNN-IV	0.05	45	–	0.01463	0.12095	0.10914
	RBFNN-V	0.005	85	–	0.00498	0.07054	0.09651
	RBFNN-VI	0.001	15	–	0.00499	0.07064	0.09210
	RBFNN-VII	0.001	150	–	0.00474	0.26617	0.19059
	RBFNN-VIII	0.001	200	–	0.00474	0.06884	0.09749
	RBFNN-IX	0.005	150	–	0.00498	0.07054	0.09652

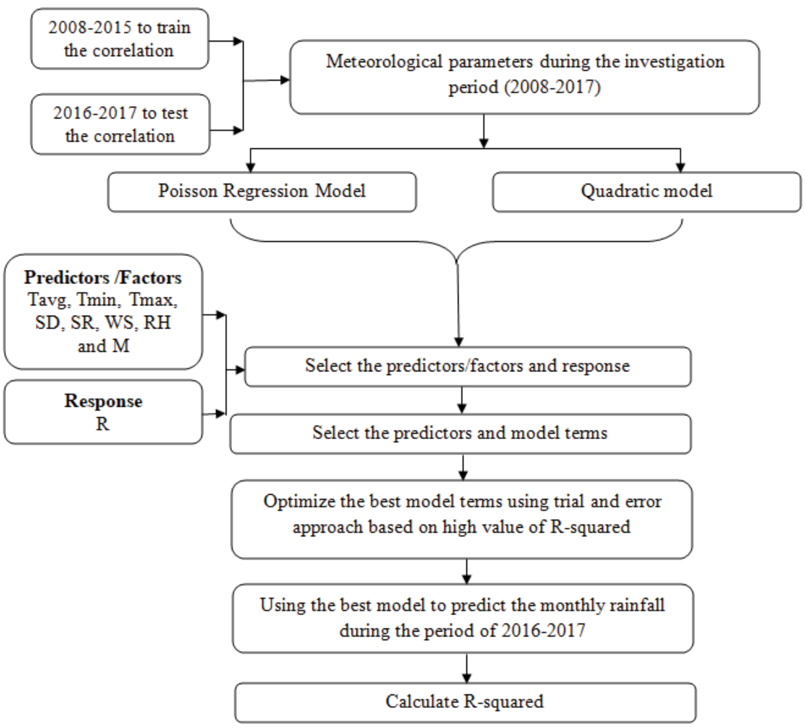


Fig. 7. Flowchart of the MM based method prediction procedure.

Table 7
Results of predicting the monthly rainfall and performance evaluation of the models

Month	Actual	MFFNN	CFNN	RBFNN	PRM	QM
112	0	25.8	30.2	14.9	22.3	3.1
105	116.4	118.6	164.1	131.0	148.3	134.5
111	0	4.0	1.2	–1.2	0.0	14.0
118	2.6	13.6	11.3	–52.6	0.9	–10.0
98	0	8.5	1.1	–44.1	0.0	–24.3
101	72	16.9	28.0	–45.3	25.3	0.8
100	9	5.8	2.3	–6.8	3.6	11.4
109	0	3.5	1.8	–90.6	0.0	–17.5
115	144.3	146.9	420.0	182.1	221.1	200.7
104	323.9	229.1	261.3	306.7	284.3	294.1
116	257	343.2	511.6	412.5	354.2	374.1
120	0	17.1	4.2	–101.5	0.0	–17.3
102	214.9	113.4	218.7	127.1	135.8	130.8
106	2	22.1	17.6	–15.9	20.2	2.5
110	0	12.8	0.5	–39.4	0.0	–5.9
119	0	7.1	5.5	–54.3	0.0	15.0
108	0	15.3	28.8	–44.4	1.6	–11.1
117	163.9	122.4	456.2	152.9	154.0	154.9
113	48.1	45.8	158.7	–63.9	58.0	3.7
103	211.7	165.6	242.8	170.7	171.4	179.5
99	1	8.1	7.0	9.5	8.0	15.1
97	0	18.4	6.0	–29.2	0.0	–5.2
107	0	20.2	6.2	–33.1	1.7	–9.9
114	184.5	175.8	324.1	157.2	204.9	184.7
R^2	–	0.849	0.714	0.835	0.887	0.879
RMSE*	–	39.009	105.940	63.069	35.069	38.521
RMSE**	–	0.0603	0.1637	0.0974	0.0542	0.0595
NSE	–	0.846	–0.136	0.597	0.875	0.850
d	–	0.955	0.829	0.922	0.969	0.965

*RMSE for non-normalized data.
**RMSE for normalized data. The limits of data are listed in Table 6.

The rainfed agricultural system is significantly impacted by rainfall in addition to anthropogenic forces. Based on the analysis, it is found that the mean values of monthly rainfall were within the range of 66.7–147.1 mm during the investigation period. In addition, it is observed that the amount of rainfall shows strong positive correlation with temperature and relative humidity. Developing an accurate model to capture the dynamic connection between rainfall and weather parameters remains a problematic task for engineers. In this study, two proposed models (QM and PRM) are used to predict monthly rainfall in Jigawa State, Nigeria, and compared with three popular machine learning algorithms (MFFNN, CFNN, and RBFNN) in order to obtain more accurate results when predicting the monthly rainfall. Based on the findings, the lowest value RMSE of 0.0542 and highest R^2 of 0.887 are provided by the PRM model, that is, the performance of PRM was better than the other models (Table 7). Therefore, the PRM model can better represent the relationship between the meteorological parameters and rainfall and produce a better prediction of the monthly rainfall. It can be concluded that the PRM model performed better than the QM and machine learning models because it was fitted based on the limited number of samples that were available in this study. To ensure the accuracy of the proposed model, the performance results of the PRM model are compared to previous scientific studies, which used meteorological parameters as input for the predictive model to predict the monthly rainfall [16,28,38]. Bagirov et al. [16] proposed the clusterwise linear regression technique for the prediction of monthly rainfall and compared it with multiple linear regression, ANNs, and the support vector machines. The results indicated that the proposed algorithm outperformed other methods in most locations based on RMSE, which ranged from 19.7 to 39.3. Anh et al. [28] introduced novel hybrid models for monthly rainfall prediction, which were combined of two pre-processing methods (seasonal decomposition and discrete wavelet transform) and two feed-forward neural networks (ANN and Seasonal ANN). The results showed that the model with the combination of Meyer wavelet and seasonal ANN provided the lowest RMSE and highest R^2

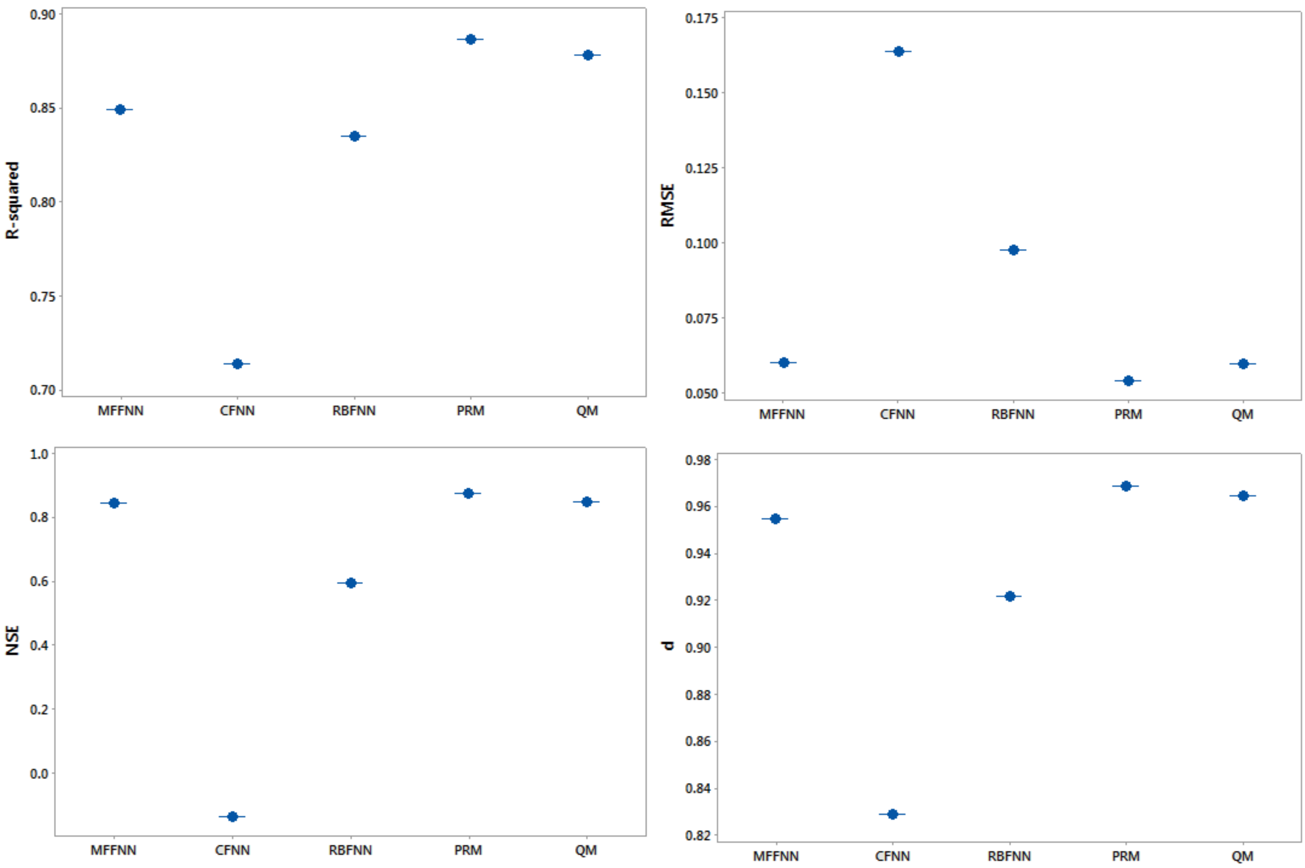


Fig. 8. Comparisons between five models used to predict monthly rainfall.

with values of 12.105 and 0.9973, respectively. Also, among the models, it was found that ARIMA model had the lowest value of R^2 (0.7628) and highest value of RMSE (108.07). Gökçekuş et al. [38] developed 25 ANN models to predict the monthly rainfall by varying the meteorological parameters. The results showed that ANN-17 with the combination of (T_{min}, T_{max}, SD, GSR) had the maximum R^2 (0.6488) compared to the other models. Additionally, based on RMSE, they found that ANN-23 with a combination of (T_{min}, T_{max}, T_{av}, W, SD) gave the lowest value of RMSE (0.1259) and was the best fit for predicting the monthly rainfall. Consequently, it was concluded that the proposed models could satisfactorily simulate non-stationary and non-linear time series-related problems such as rainfall prediction, but PRM provided the most accurate prediction for monthly rainfall.

5. Conclusions

Due to the water scarcity rainfed agriculture will continue to be the major source of food for the rapidly increasing population in Jigawa state in Nigeria, which is considered one of the most agriculturally endowed states in the country. The significant agricultural water scarcity in the country is more associated with the variability of rainfall. Therefore, this paper examined the impact of meteorological parameters including monthly average temperature, minimum temperature, maximum temperature, relative humidity, solar radiation, sunshine duration, and wind speed on monthly rainfall by utilizing correlation analysis in terms of temporal correlation and spatial correlation analysis. The results indicated that the relative humidity and temperature have a positive impact on the variability of rainfall in the selected region. Also, the monthly rainfall has been analyzed statistically and the type of distribution functions has been selected based on the skewness and kurtosis values. The results showed that the most frequent distribution at the selected region is type III, which is characterized by a narrow peak with positive till for monthly rainfall. Moreover, to enable the design of a model with accurate prediction, this paper summarized the recent scientific studies aimed at predicting the rainfall in Nigeria and around the world utilizing artificial and mathematical models. According to this review, QM and PRM have not yet been considered in other studies about monthly rainfall prediction. Therefore, to address the main objective of the current study, the authors proposed QM and PRM to predict the monthly rainfall as a function of monthly average temperature, minimum temperature, maximum temperature, relative humidity, solar radiation, sunshine duration, and wind speed. In addition, the monthly rainfall was evaluated through three artificial models, namely MFFNN, CFFNN, and RBNN based on the measurement data. The inputs of the model were monthly average temperature, minimum temperature, maximum temperature, relative humidity, solar radiation, sunshine duration, and wind speed. The proposed models were then compared in terms of predictive accuracy to select the best model. The results indicated that the developed PRM was superior in predicting the value of monthly rainfall with reported values of 0.887, 0.0542, 0.875, and 0.969 for the parameters of R^2, RMSE, NSE, and d respectively.

References

[1] F. Frank, E. Viglizzo, Water use in rain-fed farming at different scales in the Pampas of Argentina, Agric. Syst., 109 (2012) 35–42.
[2] B. Biazin, G. Sterk, M. Temesgen, A. Abdulkedir, L. Stroosnijder, Rainwater harvesting and management in rainfed agricultural systems in sub-Saharan Africa – a review, Phys. Chem. Earth Parts A/B/C, 47–48 (2012) 139–151.
[3] S. Dercon, L. Christiaensen, Consumption Risk, Technology Adoption and Poverty Traps: Evidence from Ethiopia, World Bank Policy Research Working Paper 4257, SSRN Electronic Journal, 2008.
[4] S.D. Falco, J.P. Chavas, On crop biodiversity, risk exposure, and food security in the highlands of Ethiopia, Am. J. Agric. Econ., 91 (2009) 599–611.
[5] M. Amare, N.D. Jensen, B. Shiferaw, J.D. Cissé, Rainfall shocks and agricultural productivity: implication for rural household consumption, Agric. Syst., 166 (2018) 79–89.
[6] O.E. Olayide, T. Alabi, Between rainfall and food poverty: assessing vulnerability to climate change in an agricultural economy, J. Cleaner Prod., 198 (2018) 1–10.
[7] O.E. Olayide, I.K. Tetteh, L. Popoola, Differential impacts of rainfall and irrigation on agricultural production in Nigeria: any lessons for climate-smart agriculture?, Agric. Water Manage., 178 (2016) 30–36.
[8] A.F. Ribeiro, A. Russo, C.M. Gouveia, P. Páscoa, Copula-based agricultural drought risk of rainfed cropping systems, Agric. Water Manage., 223 (2019) 105689, doi: 10.1016/j.agwat.2019.105689.
[9] National Bureau of Statistics. Contributions to Gross Domestic Products. Available at: http://nigerianstat.gov.ng (Accessed February 16, 2020)
[10] H. Fjelde, N.V. Uexkull, Climate triggers: rainfall anomalies, vulnerability and communal conflict in Sub-Saharan Africa, Political Geogr., 31 (2012) 444–453.
[11] W. Sha, K. Edwards, The use of artificial neural networks in materials science based research, Mater. Des., 28 (2007) 1747–1752.
[12] F.S. Mjalli, S. Al-Asheh, H. Alfadala, Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance, J. Environ. Manage., 83 (2007) 329–338.
[13] S.C. Keat, B.B. Chun, L.H. San, M.Z.M. Jafri, Multiple regression analysis in modelling of carbon dioxide emissions by energy consumption use in Malaysia, AIP Conf. Proc., 1657 (2015) 1–5, doi: 10.1063/1.4915185.
[14] H.D. Purnomo, K.D. Hartomo, S.Y.J. Prasetyo, Artificial neural network for monthly rainfall rate prediction, IOP Conf. Ser.: Mater. Sci. Eng., 180 (2017) 1–9, doi: 10.1088/1757-899X/180/1/012057.
[15] T.S. Abdulkadir, A.W. Salami, A.S. Aremu, A.M. Ayanshola, D.O. Oyejobi, Assessment of neural networks performance in modeling rainfall amounts, J. Res. For. Wildl. Environ., 9 (2017) 12–22.
[16] A.M. Bagirov, A. Mahmood, A. Barton, Prediction of monthly rainfall in Victoria, Australia: clusterwise linear regression approach, Atmos. Res., 188 (2017) 20–29.
[17] T. Kashiwao, K. Nakayama, S. Ando, K. Ikeda, M. Lee, A. Bahadori, A neural network-based local rainfall prediction system using meteorological data on the internet: a case study using data from the Japan Meteorological Agency, Appl. Soft Comput., 56 (2017) 317–330.
[18] Y. Xiang, L. Gou, L. He, S. Xia, W. Wang, A SVR–ANN combined model based on ensemble EMD for rainfall prediction, Appl. Soft Comput., 73 (2018) 874–883.
[19] R. Mirabbasi, O. Kisi, H. Sanikhani, S.G. Meshram, Monthly long-term rainfall estimation in Central India using M5Tree, MARS, LSSVR, ANN and GEP models, Neural Comput. Appl., 31 (2018) 6843–6862.
[20] M. Zeynoddin, H. Bonakdari, A. Azari, I. Ebtehaj, B. Gharabaghi, H.R. Madavar, Novel hybrid linear stochastic with non-linear extreme learning machine methods for forecasting monthly rainfall a tropical climate, J. Environ. Manage., 222 (2018) 190–206.
[21] A. Bello, M. Mamman, Monthly rainfall prediction using artificial neural network: a case study of Kano, Nigeria, Environ. Earth Sci. Res. J., 5 (2018) 37–41.
[22] N. Rodi, M. Malek, A. Ismail, Monthly rainfall prediction model of peninsular Malaysia using clonal selection algorithm, Int. J. Eng. Technol., 7 (2018) 182–185.
[23] S. Hudnurkar, N. Rayavarapu, Performance of Artificial Neural Network in Now Casting Summer Monsoon Rainfall: A Case Study, Conference: IEEE Punecon, Pune, 2018.
[24] E.E. Peter, E.E. Precious, Skill comparison of multiple-linear regression model and artificial neural network model in seasonal rainfall prediction-north east Nigeria, Asian Res. J. Math., 11 (2018) 1–10.
[25] S. Chattopadhyay, G. Chattopadhyay, Conjugate gradient descent learned ANN for Indian summer monsoon rainfall and efficiency assessment through Shannon-Fano coding, J. Atmos. Sol. Terr. Phys., 179 (2018) 202–205.
[26] Y. Dash, S.K. Mishra, B.K. Panigrahi, Rainfall prediction for the Kerala state of India using artificial intelligence approaches, Comput. Electr. Eng., 70 (2018) 66–73.
[27] R. Mohammadpour, Z. Asaie, M.R. Shojaeian, M. Sadeghzadeh, A hybrid of ANN and CLA to predict rainfall, Arabian J. Geosci., 11 (2018), doi: 10.1007/s12517-018-3804-z.
[28] D.T. Anh, T.D. Dang, S.P. Van, Improved rainfall prediction using combined pre-processing methods and feed-forward neural networks, J, 2 (2019) 65–83.
[29] I.R. Ilaboya, O.E. Igbinedion, Performance of multiple linear regression (MLR) and artificial neural network (ANN) for the prediction of monthly maximum rainfall in Benin City, Nigeria, Int. J. Eng. Sci. Appl., 3 (2019) 21–37.
[30] L.C.P. Velasco, R.P. Serquiña, M.S.A. Zamad, B.F. Juanico, J.C. Lomocso, Week-ahead rainfall forecasting using multilayer perceptron neural network, Procedia Comput. Sci., 161 (2019) 386–397.
[31] I. Hossain, H.M. Rasel, M. Imteaz, F. Mekanik, Long-term seasonal rainfall forecasting using linear and non-linear modelling approaches: a case study for Western Australia, Meteorol. Atmos. Phys., 132 (2019) 131–141.
[32] Y. Lin, P.C. Lee, K.C. Ma, C.C. Chiu, A hybrid grey model to forecast the annual maximum daily rainfall, KSCE J. Civ. Eng., 23 (2019) 4933–4948.
[33] A.P. Ayodele, E.E. Precious, Seasonal rainfall prediction in Lagos, Nigeria using artificial neural network, Asian J. Res. Comput. Sci., 3 (2019) 1–10.
[34] N. Bensafi, M. Lazri, S. Ameur, Novel WkNN-based technique to improve instantaneous rainfall estimation over the north of Algeria using the multispectral MSG SEVIRI imagery, J. Atmos. Sol. Terr. Phys., 183 (2019) 110–119.
[35] S.H. Pour, A.K.A. Wahab, S. Shahid, Physical-empirical models for prediction of seasonal rainfall extremes of Peninsular Malaysia, Atmos. Res., 233 (2020) 104720, doi: 10.1016/j.atmosres.2019.104720.
[36] B.T. Pham, L.M. Le, T.T. Le, K.T. Bui, V.M. Le, H.B. Ly, I. Prakash, Development of advanced artificial intelligence models for daily rainfall prediction, Atmos. Res., 237 (2020) 104845, doi: 10.1016/j.atmosres.2020.104845.
[37] M. Ali, R. Prasad, Y. Xiang, Z.M. Yaseen, Complete ensemble empirical mode decomposition hybridized with random forest and kernel ridge regression model for monthly rainfall forecasts, J. Hydrol., 584 (2020) 124647, doi: 10.1016/j.jhydrol.2020.124647.

[38] H. Gökçekuş, Y. Kassem, J. Aljamal, Analysis of different combinations of meteorological parameters in predicting rainfall with an ANN approach: a case study in Morphou, Northern Cyprus, Desal. Water Treat., 177 (2020) 350–362.

[39] L. Diop, S. Samadianfard, A. Bodian, Z.M. Yaseen, M.A. Ghorbani, H. Salimi, Annual rainfall forecasting using hybrid artificial intelligence model: integration of multilayer perceptron with whale optimization algorithm, Water Resour. Manage., 34 (2020) 733–746.

[40] K.L. Chong, S.H. Lai, Y. Yao, A.N. Ahmed, W.Z. Jaafar, A. El-Shafie, Performance enhancement model for rainfall forecasting utilizing integrated wavelet-convolutional neural network, Water Resour. Manage., 34 (2020) 2371–2387.

[41] Seo, Detailed Explanation of the Geography of Jigawa State under the following Headings: Location, Position, Size, Population, People, Climate, Vegetation, Drainage, Mineral Resources, Economic Activities and Developmental Infrastructural Activities, 2020. Available at: http://nurt9jageneral.blogspot.com/2016/10/detailed-explanation-of-geography-of.html (Accessed May 2, 2020)

[42] A.A. Mamoon, A. Rahman, Selection of the best fit probability distribution in rainfall frequency analysis for Qatar, Nat. Hazards, 86 (2016) 281–296.

[43] Y. Kassem, H. Gökçekuş, Water resources and rainfall distribution function: a case study in Lenanon, Desal. Water Treat., 177 (2020) 306–321.

[44] Y. Kassem, H. Çamur, S.M.A. Alhuoti, Solar energy technology for northern Cyprus: assessment, statistical analysis, and feasibility study, Energies, 13 (2020) 940–969, doi: 10.3390/en13040940.

[45] M. Javari, Assessment of temperature and elevation controls on spatial variability of rainfall in Iran, Atmosphere, 8 (2017) 45–85, doi: 10.3390/atmos8030045.

[46] J. Zou, Y. Han, S. So, Overview of artificial neural networks, Methods Mol. Biol., 458 (2008) 14–22.

[47] Y. Zheng, M.S. Shadloo, H. Nasiri, A. Maleki, A. Karimipour, I. Tlili, Prediction of viscosity of biodiesel blends using various artificial model and comparison with empirical correlations, Renewable Energy, 153 (2020) 1296–1306.

[48] A. Pwasong, S. Sathasivam, A new hybrid quadratic regression and cascade forward backpropagation neural network, Neurocomputing, 182 (2016) 197–209.

[49] A. Hedayat, H. Davilu, A.A. Barfrosh, K. Sepanloo, Estimation of research reactor core parameters using cascade feed forward artificial neural networks, Prog. Nucl. Energy, 51 (2009) 709–718.

[50] A. Barati-Harooni, A. Najafi-Marghmaleki, An accurate RBF-NN model for estimation of viscosity of nanofluids, J. Mol. Liq., 224 (2016) 580–588.

[51] J. Sadler, J. Goodall, M. Morsy, K. Spencer, Modeling urban coastal flood severity from crowd-sourced flood reports using Poisson regression and random forest, J. Hydrol., 559 (2018) 43–55.

[52] S. Coxe, S.G. West, L.S. Aiken, The analysis of count data: a gentle introduction to poisson regression and its alternatives, J. Personality Assess., 91 (2009) 121–136.

[53] J. Cho, J. Lee, Multiple linear regression models for predicting nonpoint-source pollutant discharge from a highland agricultural region, Water, 10 (2018) 1156–1173, doi: 10.3390/w10091156.

[54] G. Tegegne, D.K. Park, Y.O. Kim, Comparison of hydrological models for the assessment of water resources in a data-scarce region, the Upper Blue Nile River Basin, J. Hydrol.: Reg. Stud., 14 (2017) 49–66.

Supplementary information

Table S1
Results of the correlation analysis between meteorological parameters and rainfall using temporal correlation analysis

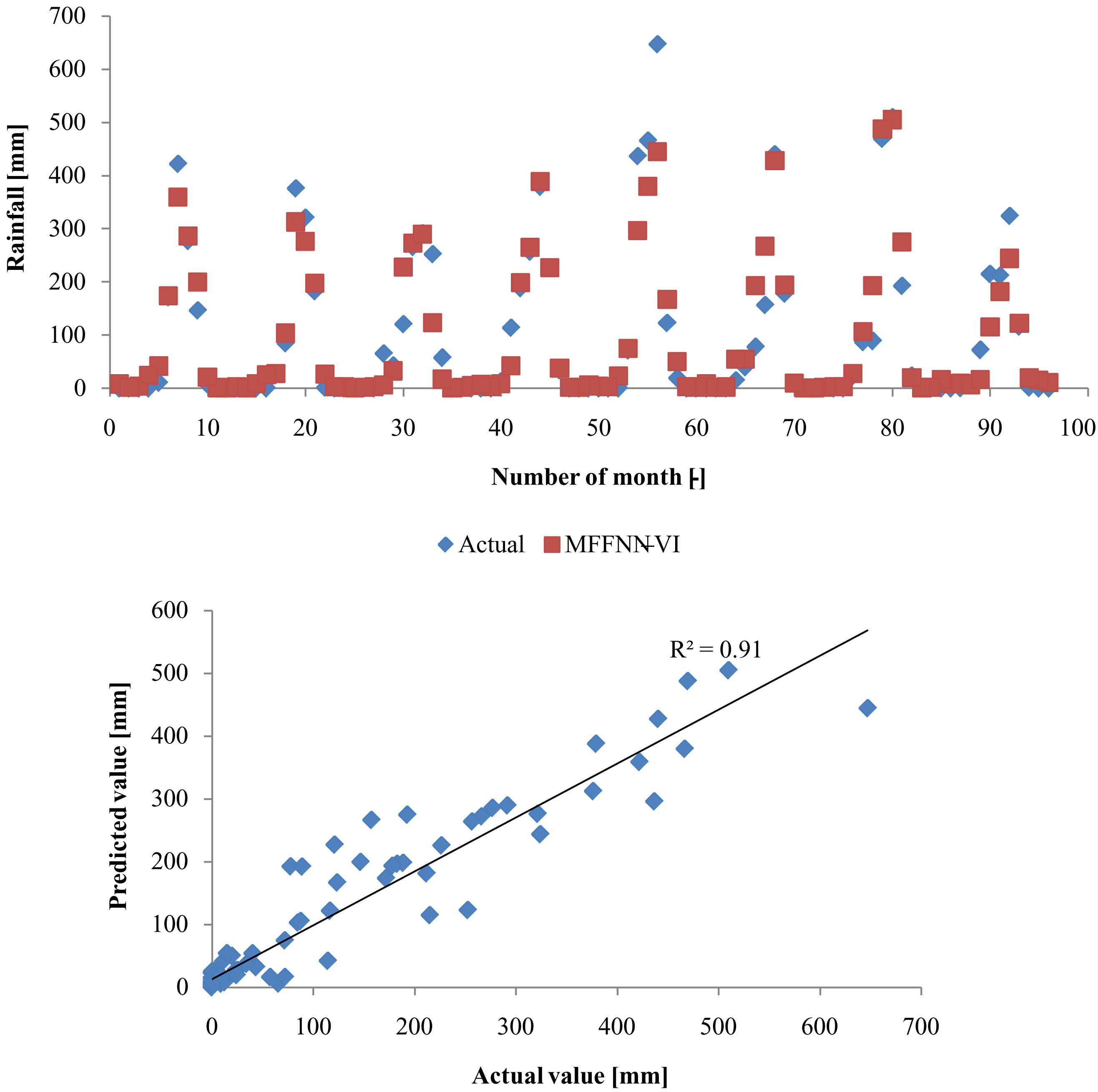
		T_{avg}	RH	SD	WS	SR	T_{min}	T_{max}	R
T_{avg}	Pearson coefficient	1							
	Significance (2-tailed)								
RH	Pearson coefficient	0.136	1						
	Significance (2-tailed)	0.140							
SD	Pearson coefficient	–0.058	–0.076	1					
	Significance (2-tailed)	0.531	0.410						
WS	Pearson coefficient	–0.031	0.099	0.083	1				
	Significance (2-tailed)	0.734	0.282	0.366					
SR	Pearson coefficient	0.188*	–0.221*	0.101	–0.083	1			
	Significance (2-tailed)	0.04	0.016	0.273	0.368				
T_{min}	Pearson coefficient	0.899**	0.434**	–0.135	–0.071	0.174	1		
	Significance (2-tailed)	0.000	0.000	0.142	0.441	0.058			
T_{max}	Pearson coefficient	0.831**	–0.282**	0.570	0.028	0.150	0.503**	1	
	Significance (2-tailed)	0.000	0.000	0.535	0.759	0.102	0.000		
R	Pearson coefficient	0.016	0.16	–0.179	0.170	–0.087	0.307**	0.359**	1
	Significance (2-tailed)	0.866	0.866	0.051	0.063	0.347	0.001	0.000	

*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).

Table S2
Results of the correlation analysis between meteorological parameters and rainfall using spatial correlation analysis

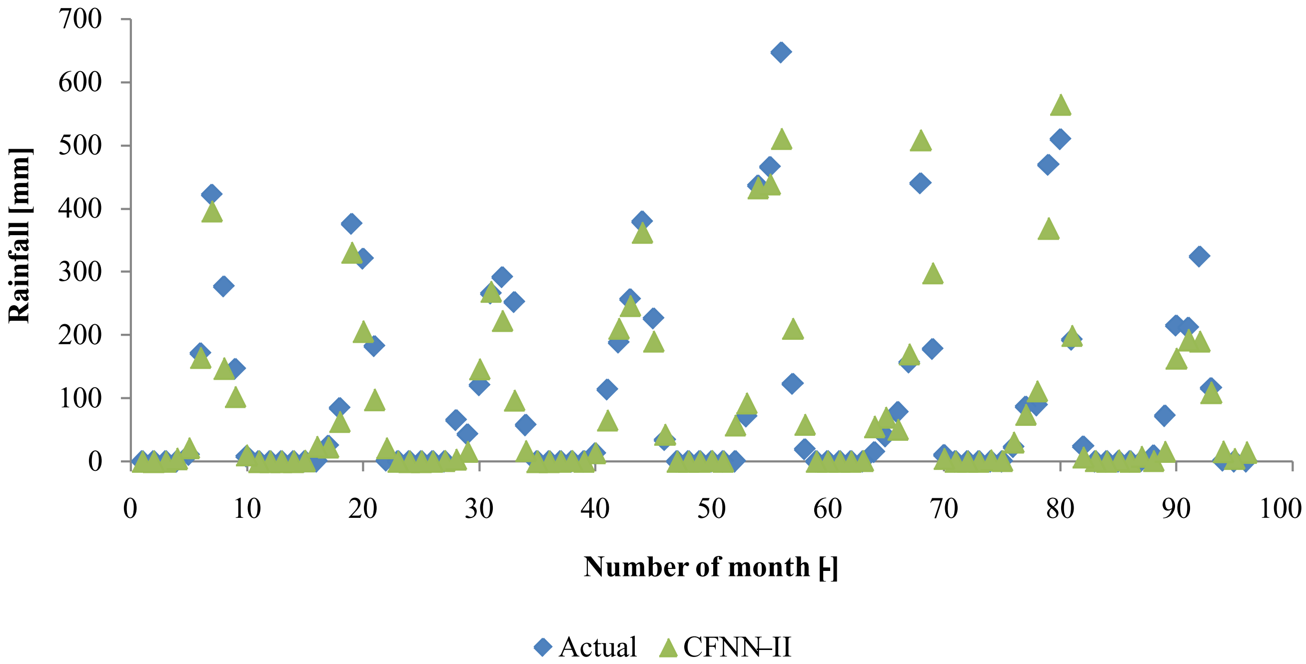
		T_{avg}	RH	SD	WS	SR	T_{min}	T_{max}	R
T_{avg}	Correlation coefficient	1							
	Significance (2-tailed)								
RH	Correlation coefficient	0.069	1						
	Significance (2-tailed)	0.453							
SD	Correlation coefficient	–0.072	–0.120	1					
	Significance (2-tailed)	0.431	0.191						
WS	Correlation coefficient	0.29	0.111	0.026	1				
	Significance (2-tailed)	0.75	0.227	0.776					
SR	Correlation coefficient	0.218*	–0.130	0.113	0.021	1			
	Significance (2-tailed)	0.017	0.157	0.221	0.820				
T_{min}	Correlation coefficient	0.886**	0.297*	–0.072	0.071	0.216*	1		
	Significance (2-tailed)	0.000	0.001	0.433	0.440	0.018			
T_{max}	Correlation coefficient	0.835**	–0.268**	0.08	0.014	0.116	0.560**	1	
	Significance (2-tailed)	0.000	0.003	0.384	0.883	0.208	0.000		
R	Correlation coefficient	0.237**	0.872**	–0.159	0.146	0.022	0.451**	–0.132	1
	Significance (2-tailed)	0.009	0.000	0.084	0.111	0.813	0.000	0.150	

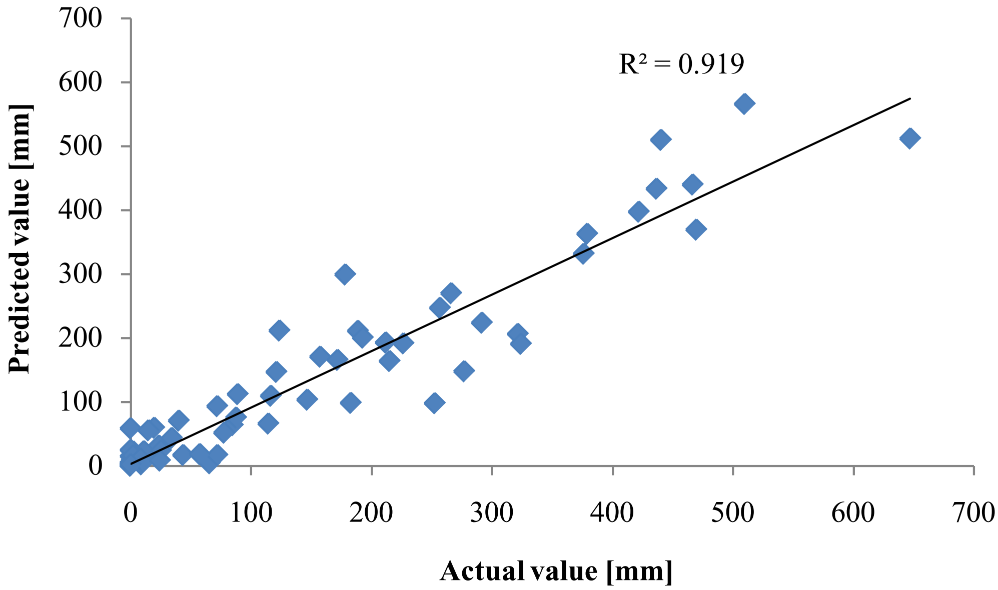
*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).



(a)

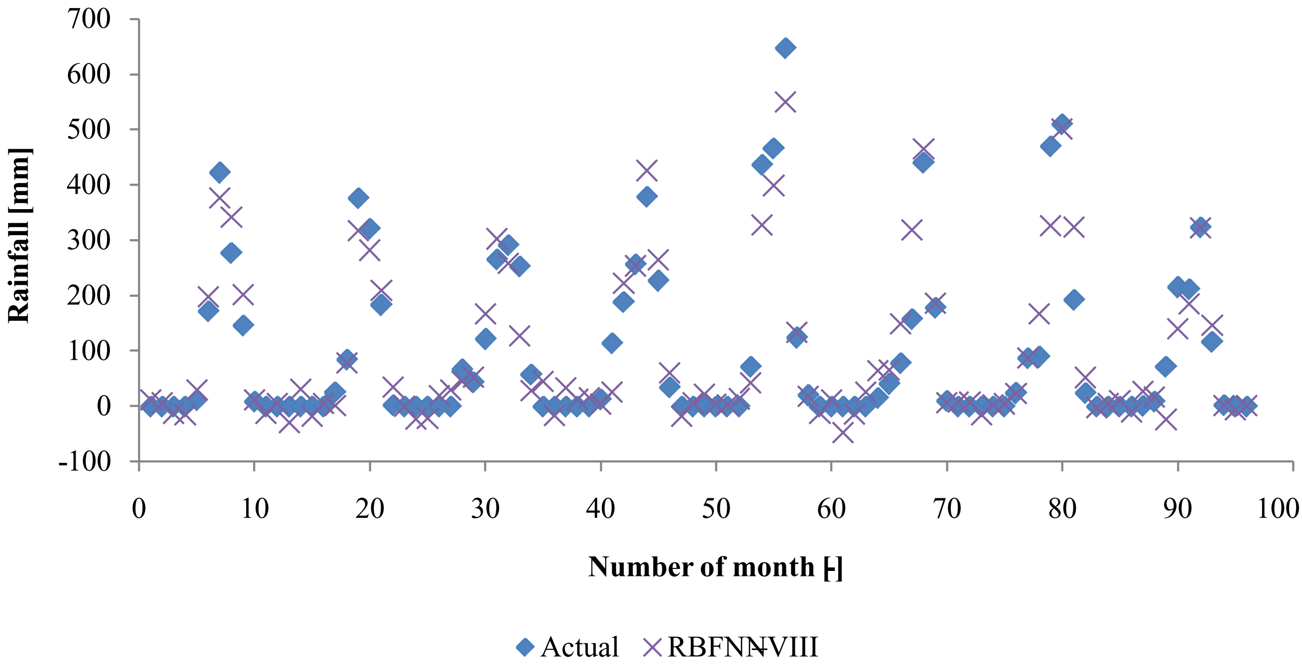
Fig. S1. Continued

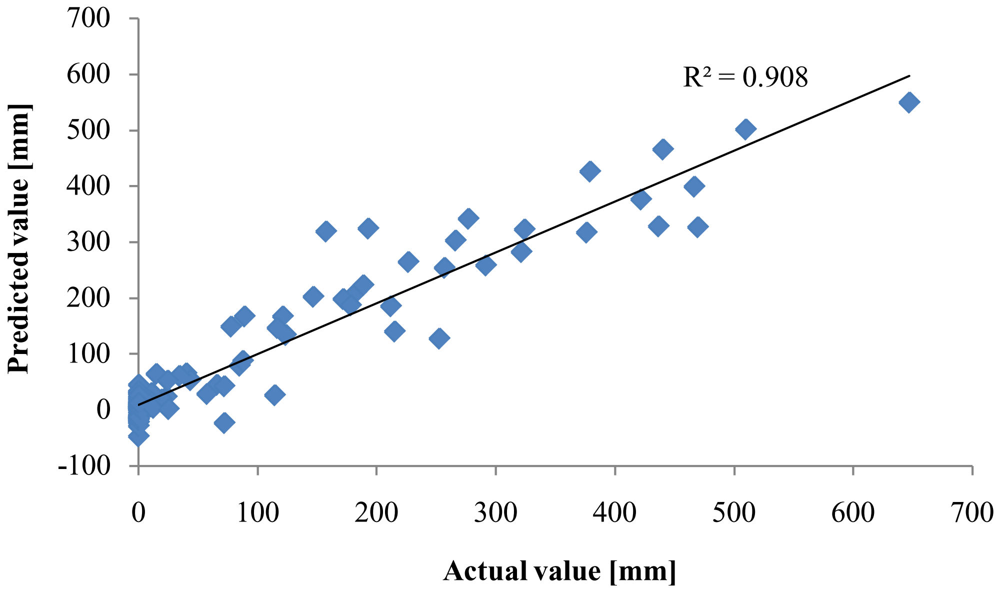




(b)

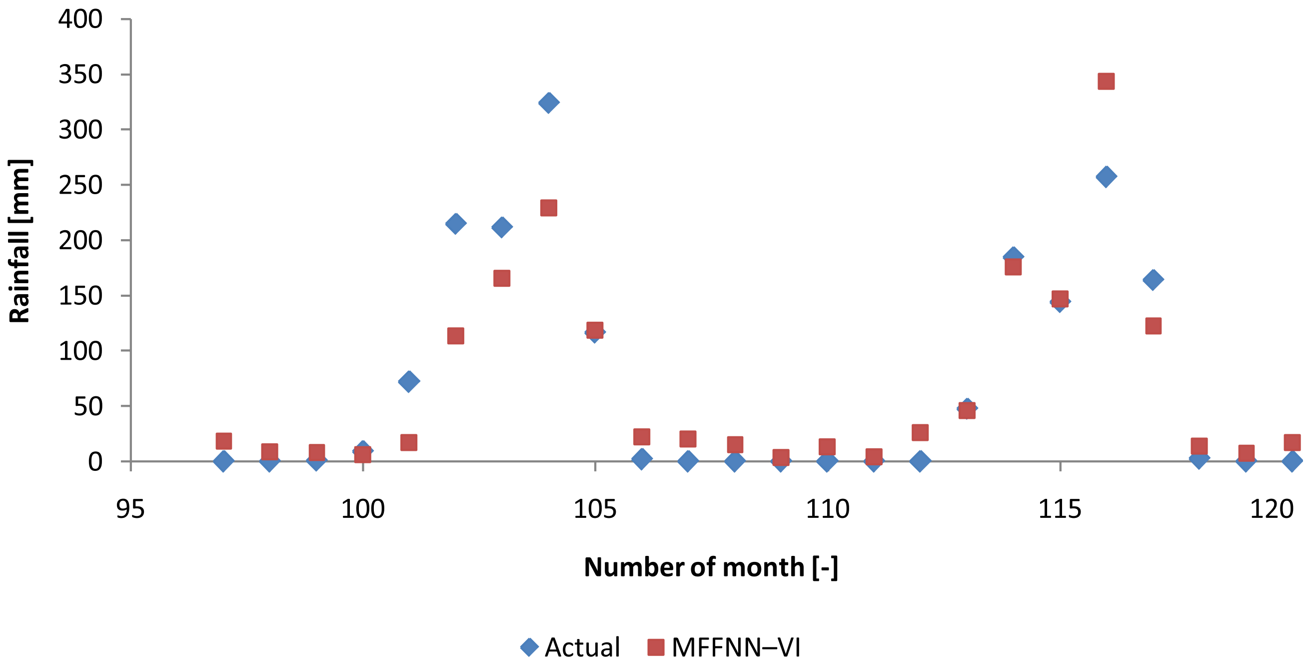
Fig. S1. Continued





(c)

Fig. S1. Comparison between actual and predicted values obtained artificial training models (a) MFFNN-VI, (b) CFNN-II, and (c) RBFNN-VIII.



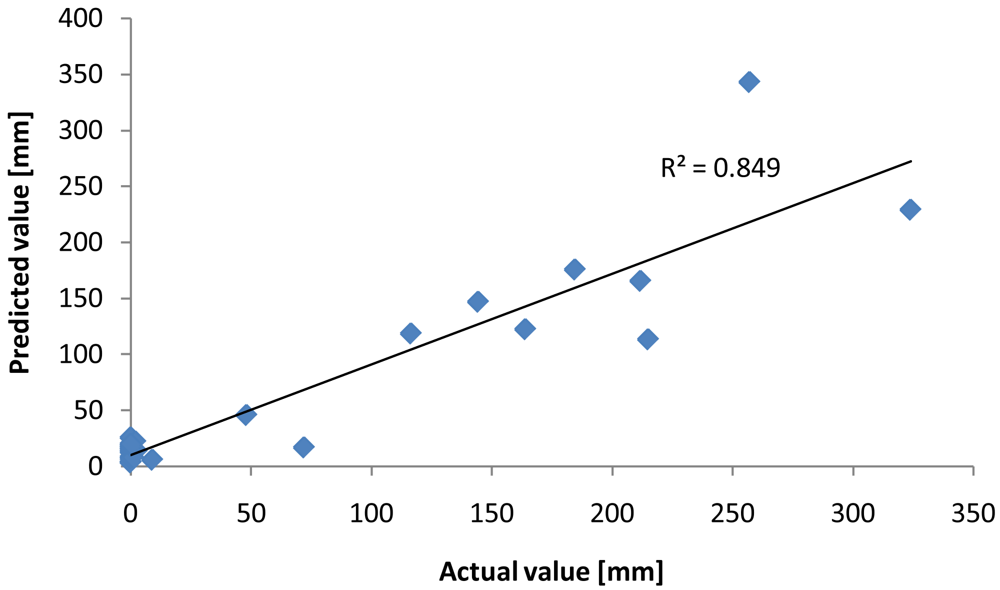


Fig. S2. Comparison between actual and predicted values obtained from MFFNN-VI.

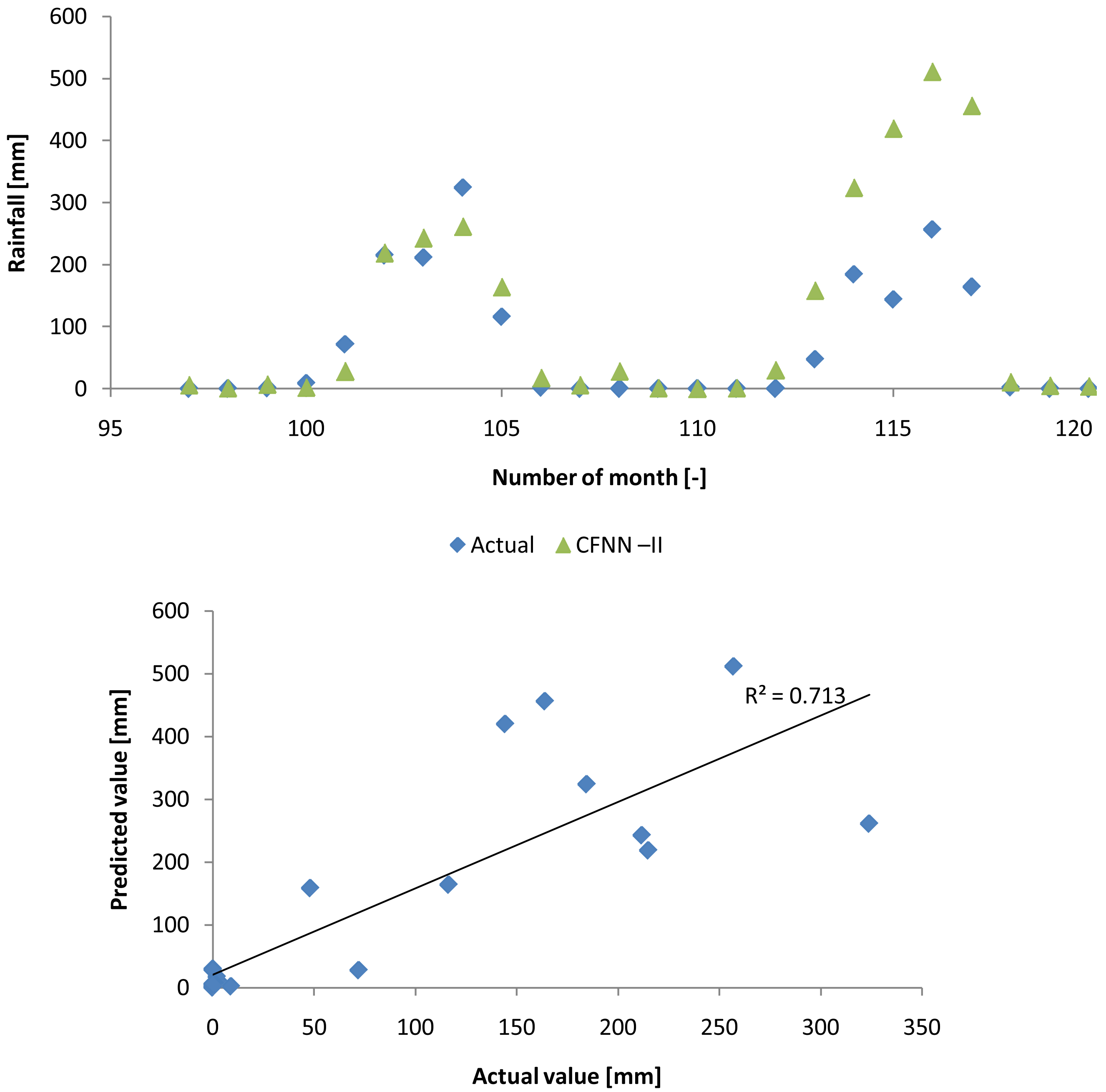


Fig. S3. Comparison between actual and predicted values obtained from CFNN-II.

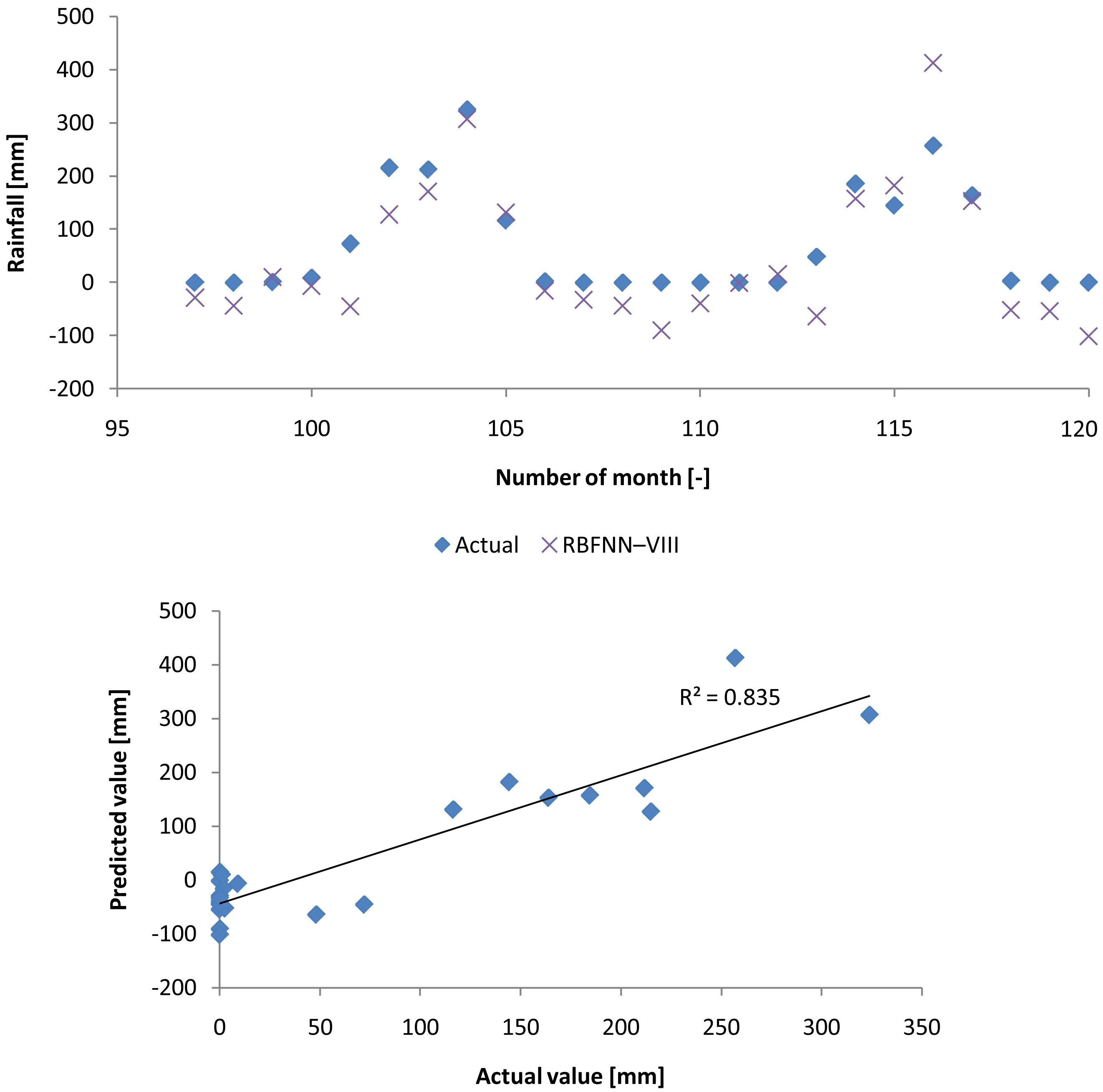
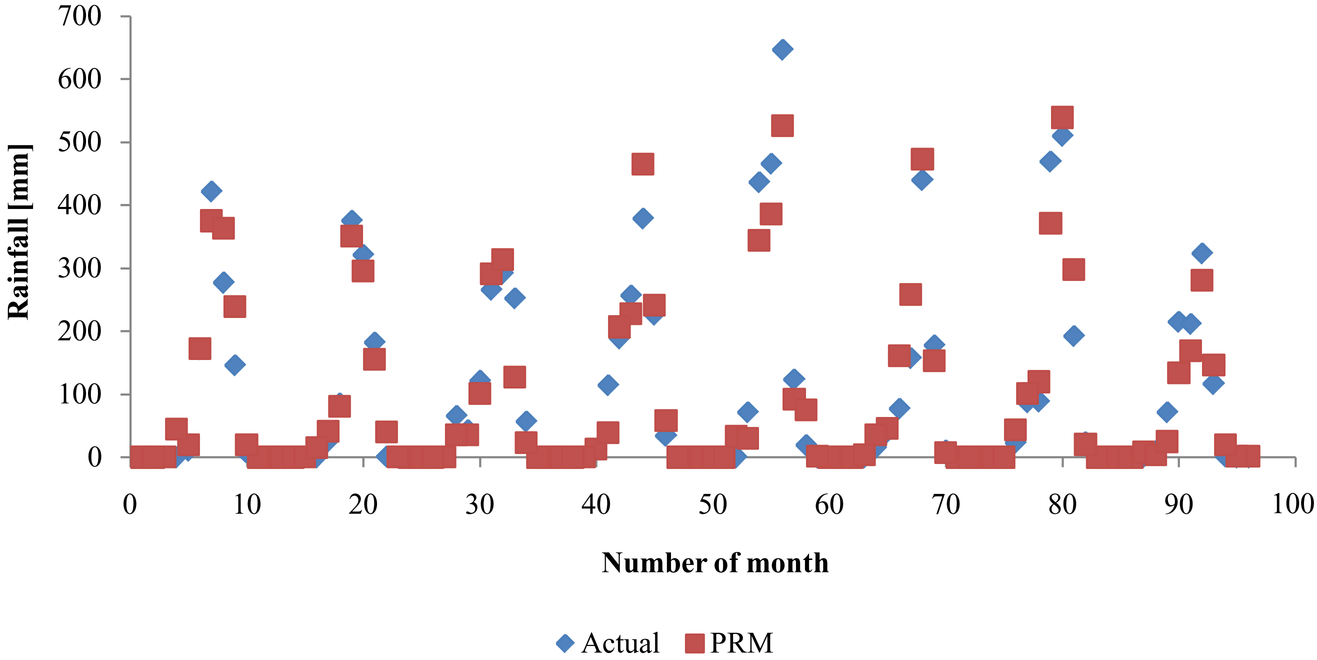
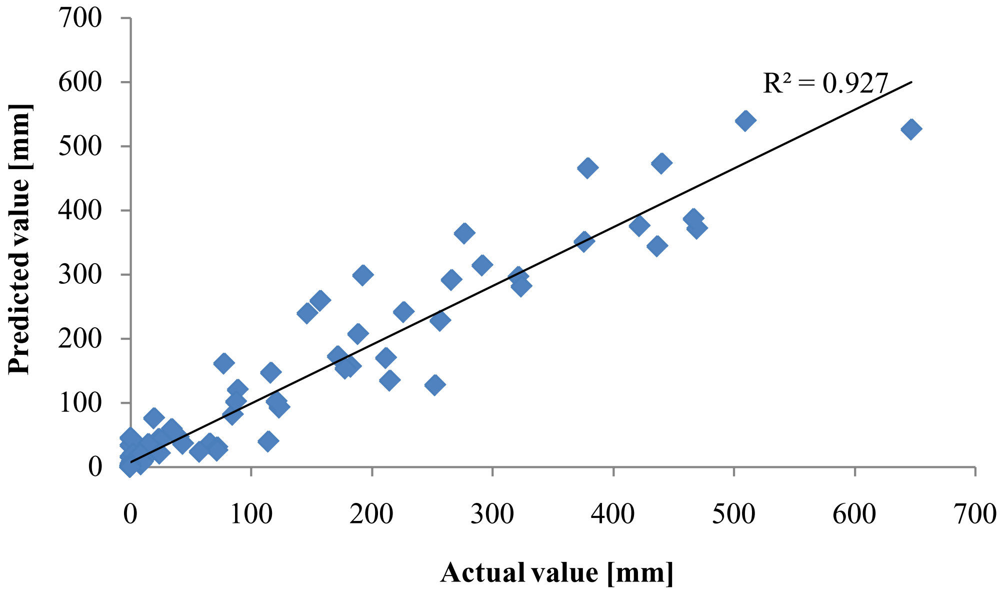


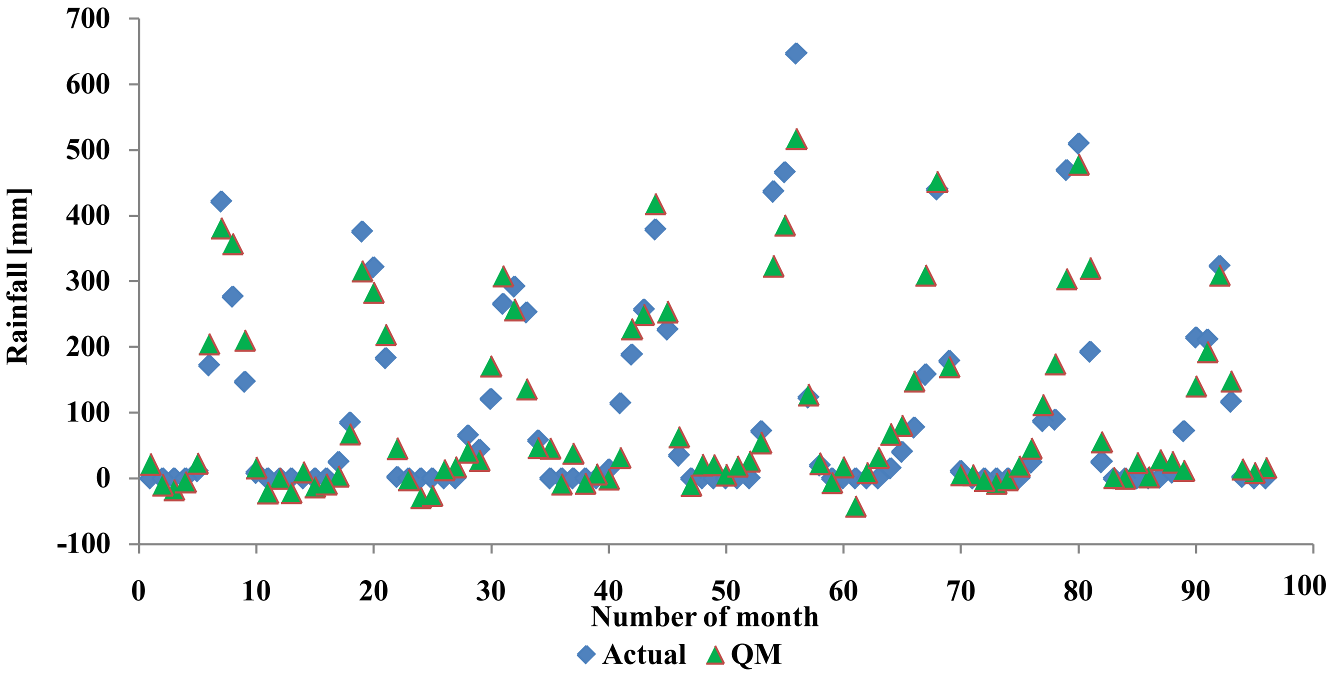
Fig. S4. Comparison between actual and predicted values obtained from RBFNN-VIII.

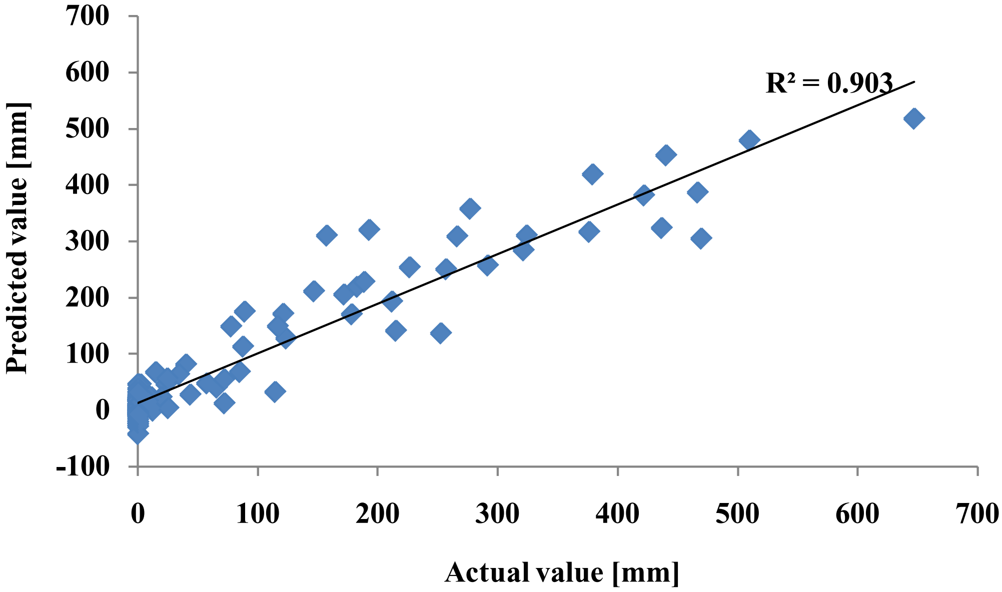




(a)

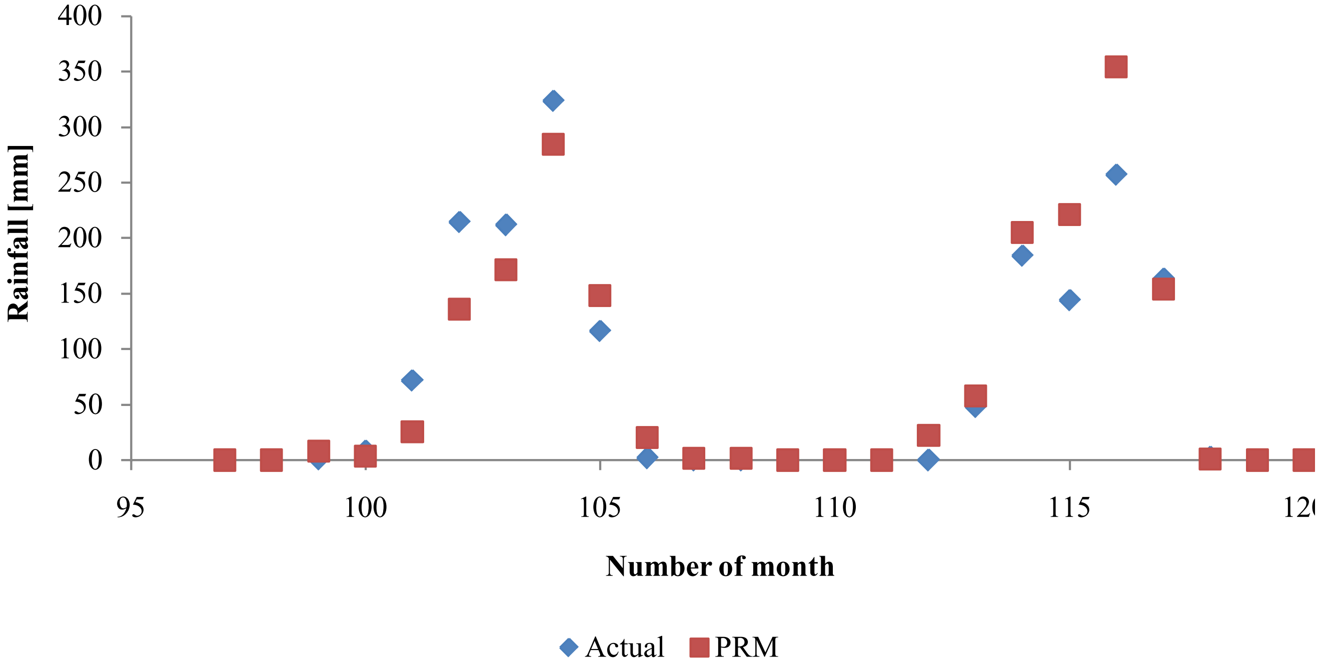
Fig. S5. Continued

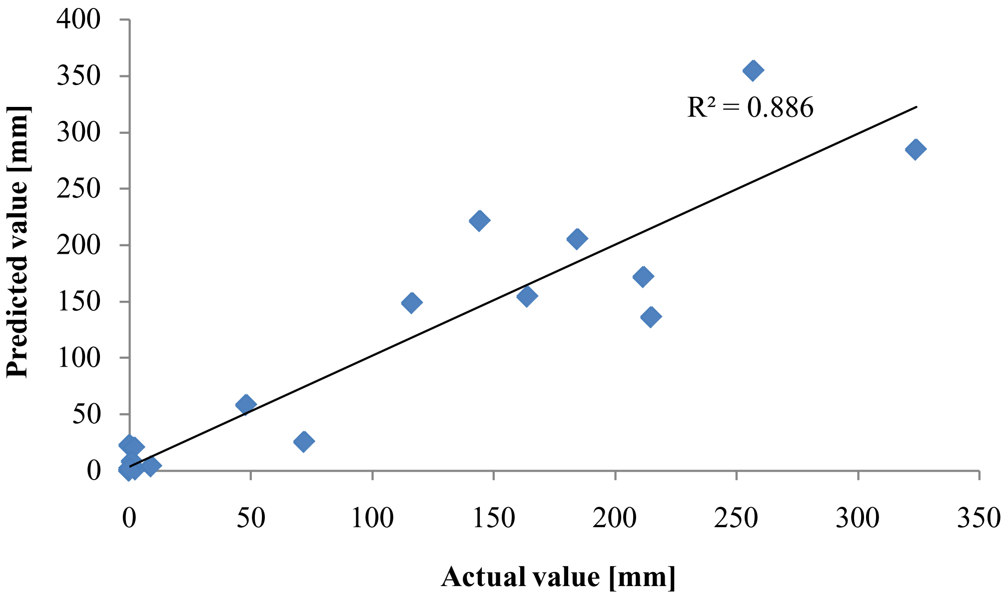




(b)

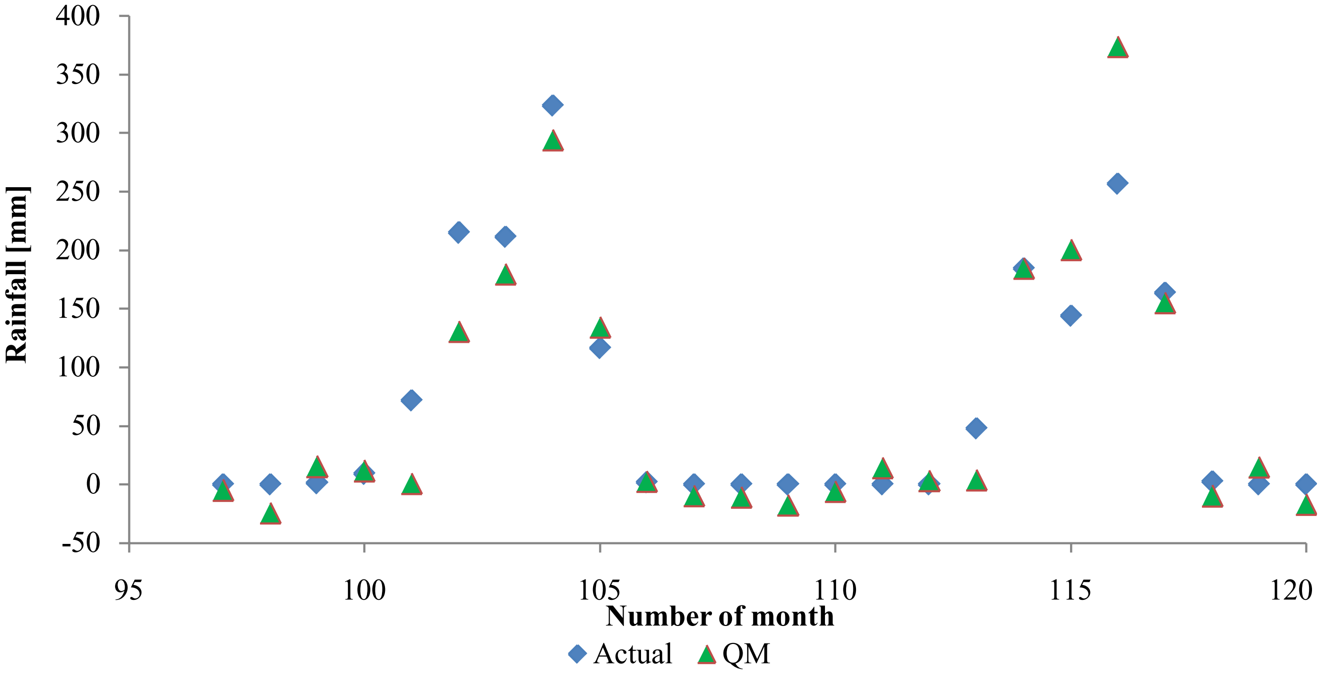
Fig. S5. Comparison between actual and predicted values obtained for training data (a) PRM model and (b) QM model.

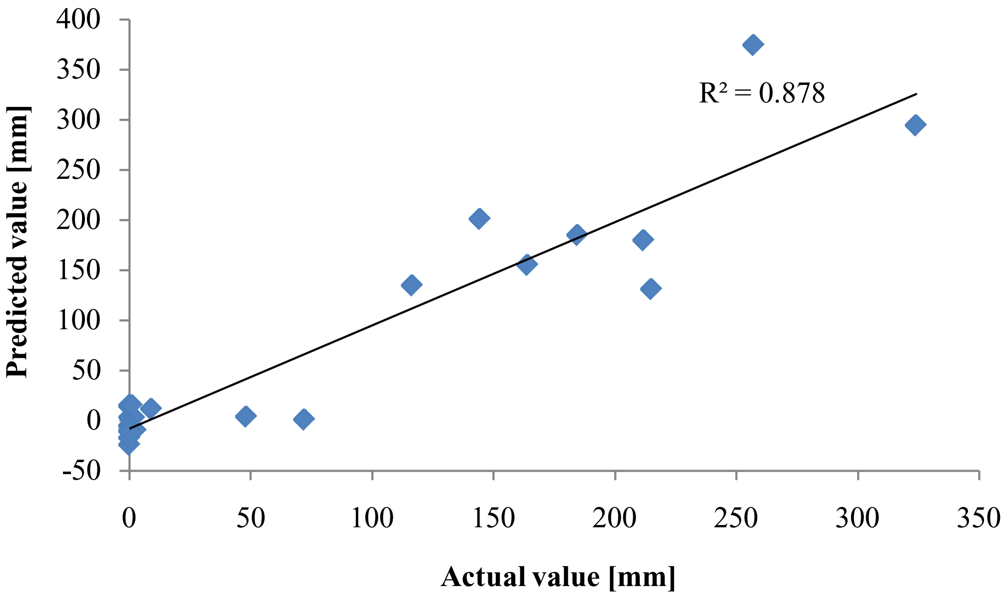




(a)

Fig. S6. Continued





(b)

Fig. S6. Comparison between actual and predicted values obtained for testing data (a) PRM model and (b) QM model.