Radial basis function neural networks for reliably forecasting rainfall


ABSTRACT
Rainfall forecasting is an interesting task especially in a modern city facing the problem of global warming; in addition rainfall is a necessary input for the analysis and design of hydrologic systems. Most rainfall real-time forecasting models are based on conceptual models simulating the complex hydrological process under climate variability. As there are a lot of variables and parameters with uncertainties and non-linear relationships, the calibration of conceptual or physically based models is often a difficult and time-consuming procedure. Simpler artificial neural network (ANN) forecasts may therefore seem attractive as an alternative model. The present research demonstrates the application of the radial basis function neural network (RBFNN) to rainfall forecasting for Alexandria City, Egypt. A significant feature of the input construction of the RBF network is based on the use of the average 10 year rainfall in each decade to forecast the next year. The results show the capability of the RBF network in forecasting the yearly rainfall and two highest rainfall monsoon months, January and December, compared with other statistical models. Based on these results, the use of the RBF model can be recommended as a viable alternative for forecasting the rainfall based on historical rainfall recorded data.

Key words | Alexandria – Egypt, artificial neural network, radial basis function, rainfall forecasting

INTRODUCTION
Rainfall is produced by a complex combination of dynamic, thermodynamic, and cloud microphysical processes operating over a large extent of space and time. Because of both the space-time variable and the complex dynamics, rainfall is a difficult quantity to measure, forecast and model. In this context, researchers investigate several alternative methods to predict the rainfall by simulating the complex hydrological processes under different climate variability. At present with the increasing demand on water besides the change in climate, controlling water resource systems is a challenging task for engineers. An accurate rainfall forecast helps to obtain optimal water distribution within technical and economic constraints in the area. There are many circumstances where empirical linear or physical-based watershed models are unsuccessful in providing accurate forecasting. Although the traditional linear models are simple and easy to implement, they are unsuitable for forecasting purposes if the underlying mechanism is nonlinear. On the other hand, because of the difficulties in formulating reasonable nonlinear models for forecasting, engineers have recently resorted to a machine-learning approach called artificial neural networks (ANN). This can be defined as a massively parallel distributed processor which has a natural propensity for storing experiential knowledge and making it available for use (Haykin 1998).

A neural network is characterized by its architecture, which represents the pattern of connection between nodes, its method of determining the connection weights, and the activation function (Fausett 1994). Caudill presented a comprehensive description of neural networks in a series of
papers (Caudill 1987, 1988, 1989). A typical ANN consists of a number of nodes that are organized according to particular arrangement. Finally, an attractive feature of ANNs is their ability to extract the relation between the inputs and outputs of a process, without the physics being clearly provided to them. They are able to provide a mapping from one multivariate space to another, given a set of data representing that mapping. Even if the data is noisy and contaminated with errors, ANNs have been known to identify the underlying rule. These properties suggest that ANNs may be well-suited to the problems of estimation and prediction in hydrology.

Many different rainfall forecasting techniques have been proposed. Since the early 1970s, researchers have used statistical model identification to forecast (Akaike 1974), although forecast development improved with the use of Markov techniques (Baum 1972; Rabiner & Juang 1986). Recently, the relationship of rainfall precipitation to synoptic atmospheric patterns has led to the development of the weather-state model (Hay et al. 1991; Bardossy & Plate 1992; Hughes et al. 1993).

Despite the long history of development and evaluation of techniques for forecasting by using artificial neural networks, forecasting accuracy is still relatively poor. Researchers worked on rainfall in two parts: first by forecasting the rainfall–runoff; second by forecasting precipitation over an area. The relationship between rainfall and runoff has been looked into by Lin & Chen (2004). The latter found that a nonlinear model illustrating the said relationship can be developed by using the radial basis function (RBF) network. The methodology is applied to an actual reservoir watershed to predict one to three readings of inflow. The Fei-Tsui Reservoir Watershed, in Northern Taiwan, was used to obtain real-time streamflows. Nine typhoons were recorded; the data was divided into three categories: training, validation, and testing. The RBF network was set up so that the input layer consisted of 21 nodes while the output layer consisted of one node only. Three of the 21 nodes were meant to represent inflow rates for the past two hours, past one hour and the present time. The remaining 18 input nodes were used to record the past two hour, past one hour, and present rainfall depths of six rain gauges. On the other hand, the output node represents the data for either one hour or two hours or three hours forecast of streamflow. Therefore, the results obtained are evidence to the fact that the RBF network can be successfully applied to determine the relationship between rainfall and runoff. Senthil Kumar et al. (2004) used the RBF to model rainfall–runoff and compared the RBF with multi-layer perception (MLP) networks; the authors studied the Malaprabha catchment in India. the RBF model has an accurate coefficient of correlation near to 1. In order to evaluate the statistical significance of the inferences drawn from this study, the author performed a similar analysis on one more basin, the Baitarani basin in India. Rainfall and runoff data for 3 years (monsoon season), 1980 to 1982, were used. The first 2 years of data were employed for training the network and the balance of the data used for validation. A significant observation is that the MLP failed to match the peak flow during validation, compared with RBF, with a −61.05% for MLP and 5.14% for RBF. The results of the study suggest that the RBF networks are a viable alternative to the already popular MLP networks.

Lee et al. (2009) used artificial neural networks to predict regional runoff utilization, using two different types of artificial neural network model (RBF and BPNN) to build up small area rainfall–runoff supply systems. A historical rainfall for a Taipei City in Taiwan was applied in the study. As a result of the impact variances between the results used in training, testing and prediction and the actual results, the overall success rates of prediction are about 85% for BPNN and 98.6% for RBF.

Jareanpon et al. (2004) used the adaptive radial basis function network (RBFN) to forecast rainfall in Thailand. A significant feature of the adaptive RBFN is that it is able to create new hidden units and solve the spread factor problem using a genetic algorithm. The author used data for 30 years (1971–2000) of monthly rainfall values at 12 climate stations. For training the network, data for the period between 1971 and 1999 was used, and for testing, data for the year 2000. Parida & Moalafhi (2008) discussed the regional rainfall frequency using L-Moments and radial bases function for Botswana. This study was based on long-term annual rainfall data from 1961 to 2003 at 11 synoptic stations across Botswana. The author concluded that RBF could be used to identify the homogeneity index through L-Moments. The goal of this project was to build an ANN model to forecast rainfall; this goal was achieved through the following objectives: develop ANN models to
predict rainfall by using radial basis function neural network type depending on rainfall data collected in the case study, and an alternative model by using multiple layer regression as a model to compare the RBFN for forecasting rainfall.

Finally, ANNs have been used in different modeling. Recently researchers started work on intelligent systems in wider environmental modeling: stream flow modeling (Fernando & Shamseldin 2009); water quality and water table (Yang et al. 1997); design of coastal sewage system (Sanchez et al. 1998); suspended sediment modeling (Rai & Mathur 2008); and optimal design of a reservoir operation system (Chandramouli & Deka 2005).

The main aim of this paper is to address rainfall in Alexandria. This study contributes to understanding, development, and implementation of an ANN model that can be used to evaluate the rainfall in future by using radial basis function neural networks (RBFNNs). The data on precipitation in Alexandria are available from 1960 to 2008 as well as the temperature and wind speed.

**METHODOLOGY**

**Study area and data**

The modeling work was carried out using the previous rainfall records of Alexandria City, Egypt, as shown in Figure 1 (a) and (b), also illustrating the location of rain gauge stations. Alexandria lies between 31°12’0” N, and 29°55’0” E. The city has a waterfront that extends for 60 km, from Abu-Qir bay in the east to Sid-Krier in the west and includes a number of beaches and harbors. The area is characterized by the irregular hills in the southern parts with an elevation from 0 to 40 meters above mean sea level and slopes towards the Mediterranean Sea in the

![Figure 1](https://iwaponline.com/jwcc/article-pdf/3/2/125/375125/125.pdf)
north. The climate of the study area is generally dry, but the prevailing north wind blowing across the Mediterranean gives the city a different climate from the desert hinterland during the monsoon months (December to March). January and February are the coolest months. Alexandria has only two seasons, a mild winter from November to April and a hot summer from May to October. The difference between the seasons is the variation in daytime temperature and changes in the prevailing wind. Average annual temperature ranges between a minimum of 14 °C in winter and maximum of 30 °C in summer. Alexandria is one of the wettest areas of Egypt, with an average annual precipitation of about 200 mm, compared with the country's annual average precipitation rate of 80 mm. Most rain falls along the coastal area and it decreases suddenly moving southwards. The humidity in Alexandria is very high; however, the sea breeze keeps the moisture down to a comfortable level. The wind velocity ranges from 30 to 38 km/h. The original data set consisted of 50 years from 1960 to 2009 for daily precipitation recorded from five rain gauges as shown in Figure 1. As the data used are from a large catchment and for daily precipitation, the analysis was done in two ways. First, we calculated the monthly rainfall by summing the daily rainfall for each month in every year. This step filters the daily rainfall data so that we can find the total yearly rainfall. Second, we evolve the monthly rainfall percentage in every year; this will predict the rainfall values during the monsoon season. The final yearly data are illustrated in Figure 2.

### The network architecture

The common multilayer feedforward network with one hidden layer of RBFN is considered for the analysis. In the RBFN in this research, the input quantities are fed to input nodes, which in turn pass them on to the hidden layer nodes which add up the weighted input received from each layer after each neuron consisting of a RBF centered on a point with as many dimensions as there are predictor variables. The structure of the RBF consists of an input layer, one hidden layer and an output layer. Figure 3 shows the structure.

In Figure 3, the number of input nodes corresponds to the number of variables in the input vector used to forecast future values. However, currently there is no suggested systematic way to determine this number. The selection of this parameter should be included in the model construction process. The hidden layer has a variable number of neurons (the optimal number is determined by the training process).
Each neuron consists of a RBF centered on a point with as many dimensions as there are predictor variables.

The RBF output layer results in a linear fashion, the output \( y \) is computed by Equation (1):

\[
y_i(x) = \sum_{k=1}^{J} W_{ki} \phi(\|X - C_k\|) \tag{1}
\]

for \( i = 1, \ldots, J \) where \( y_i(x) \) is the \( i \)th output of the RBFN, \( W_{ki} \) is the connection weight from the \( k \)th hidden unit to the \( i \)th output unit, \( C_k \) is the prototype or center of the \( k \)th hidden unit, and \( \|\cdot\| \) denotes the Euclidean norm. The RBF \( \phi(\cdot) \) is typically selected as the Gaussian function.

To calculate the center of the radial a Gaussian function has been used in each hidden unit and it depends on the distance of input from the center; the center and the spread are the parameters to be determined. It can be concluded from the Gaussian radial function that a hidden unit is more sensitive to data points near the center. The Gaussian function is described in Equation (2):

\[
\phi(x - c_i) = e^{-\|x - c_i\|^2/2\sigma^2} \tag{2}
\]

where \( c_i = (c_{i1}, c_{i2}, \ldots, c_{im}) \) is the center of the associated field, and \( \sigma \) is the width of the Gaussian function.

Finally, there is the summation layer in which the value coming out of a neuron in the hidden layer is multiplied by a weight associated with the neuron and passed to the summation layer which adds up the weighted values and presents this sum as the output of the network. Figure 4 show the three layer stages.

The RBF network in this research used supervised application and supervised learning (Chen et al. 1990, 1991). In supervised application we are provided with a set of data samples known as a training set for which the corresponding network outputs are known. The training process calculates the number of neurons in the hidden layer, the coordinate of the center of each hidden layer RBF function, the radius spread of each RBF function in each dimension, and the weights applied to the RBF function as they are passed to the summation layer. The process of learning is shown in the flowchart in Figure 5 where \( \text{num}_{\text{unit}} \) is the number of the units in hidden layer.

**Model construction for yearly and monthly forecasting**

The RBFN type models described above were first applied to predict yearly rainfall in Alexandria city. Daily rainfall in the city collected from four rainfall stations were used in this study. In general, a next year forecasting was found based on the historical data collected before. Preliminary statistical
examination of the rainfall data revealed that there are rainfall changes in each decade. It has been reported that the average rainfall changes every 10 years. Considering this phenomenon, the next year prediction model was developed by taking the average 10 year rainfall data for the city. Based on the above examination and results extended from the statistical analysis of the historical rainfall data for Alexandria city, the model construction of the RBF network is configured to consider the fact of average 10 year rainfall in each decade to find the next year. In this context, the data set of the RBF network can be expressed as:

\[ R_f(t) = f(R_y \text{ average}) \]  

(3)

where, \( R_y \) is the actual yearly rainfall and \( R_f \) is the forecasted rainfall at year (t).

To calculate the data input set in the neural network, an average for each decade was calculated as illustrated in Equation (4):

\[ R_{\text{av}}^{\text{Din}} = \frac{\sum_{i=1}^{n} R_{y_{i-1}}}{n-1} \]  

(4)

where \( Din = \) year decade, for the data here five pairs when \( i = 1 \) refer to years 1960–1969 to \( i = 5 \) refers to years 2000–2009; \( n = \) the years in each decade so, \( 1 \leq n \leq 9 \) (note for each first year in the decade it will be the total average for the 10 years before).

In order to complete the data set used in the neural network, the data need to be normalized. A simple normalization was used (Lachtermacher & Fuller 1995):

\[ x_n = x_0/x_{\text{max}} \]  

(5)

where \( x_n \) and \( x_0 \) represent the normalized and original data, \( x_{\text{max}} \) is the maximum.

The final stage in the building of the RBF neural network is the training and test sample. The training sample is used for ANN model development and the test sample is adopted for evaluating the forecasting ability of the model. Here, in yearly forecasting, the model is trained using historical rainfall data from 1960 to 2001, and tested with the rest of the data from 2002 to 2009 in order to examine the model performance and reliability. The performance measure for this network is calculated by finding the error (\( E \)) between the actual data with predicted data from the network as defined in Equation (6).

\[ E = \frac{R_p - R_s}{R_s} \times 100\% \]  

(6)

The choice of the network architecture is based on reaching the minimum error during training.

Similarly to the yearly network type, the network models were applied to predict monthly rainfall for Alexandria City. In general, a next month forecasting was found based on the historical data collected before. Preliminary statistical examination of the rainfall data revealed that the monsoon occurs over a 5-month period in Alexandria; in this research two months have been studied (January and December). It has been reported that the rainfall changes each 10 years around an average rainfall. Considering this phenomenon, the next month prediction model was developed by taking the average 10 year rainfall data for the city. The model construction of the RBF network is configured to consider the fact of average 10 year rainfall in each decade to find the next year. To calculate the data input set in the neural network, an average for each decade was calculated as shown above in the yearly stage Equation (3) but here for January and December. In order to complete the data set used in neural network, the data need to be normalized. A simple normalization was used (Lachtermacher & Fuller 1995), as shown in Equation (5). The final stage in the building of the RBF neural network is the training and test sample. The training sample is used for ANN model development and the test sample is adopted for evaluating the forecasting ability of the model. Here, in monthly forecasting, the model was trained using historical rainfall data from 1960 to 2001, and tested with the rest of the data from 2002 to 2009 for both January and December in order to examine the model performance and reliability. The performance measure for this network was calculated by finding the error (\( E \)) between the actual data with predicted data from the network as defined in Equation (6). The yearly and monthly forecasting model network architectures are shown in Figure 6.

By carrying out training of the network, the supervised learning in the network will calculate the number of neurons in the hidden layer, the coordinates of the center of each
hidden layer RBF function, the radius of each RBF function in each dimension, and the weights applied to the RBF function as they are passed to the summation layer. The training session is stopped once the minimum error is reached. In a summary for the RBF network program code, Figures 7 and 8 illustrate the training and testing procedure of the RBF code used for forecasting the rainfall in this research.

Comparison with multiple linear regressions

A statistical model of multiple regressions was employed to estimate the rainfall in Alexandria City for the same data used in the artificial neural network. This will be an alternative statistical method for forecasting the yearly rainfall and monthly rainfall.

The multiple regression models express the value of a predicted variable as a linear function of one or more
predictor variables and an error term (Holder 1985):

\[ y = a + b_1x_1 + b_2x_2 + \cdots + b_ix_i + e \] (7)

where \( y \) is a dependent variable and \( x_1, x_2, \ldots, x_i \) are independent variables, \( a, b_i \) = constant; \( e \) = random variable.

For any \( i \)th set of observations, the model can be written more conveniently as

\[ y_i = a + \beta_1(x_{i1} - \bar{x}_1) + \beta_2(x_{i2} - \bar{x}_2) + \cdots + \beta_k(x_{ik} - \bar{x}_k) + e_i \] (8)

where \( x_{ki} = \) value of independent variable \( x_k \) at \( i \)th set of observations totaling \( n \); and

\[ x_k = \frac{1}{n} \sum_{i=1}^{n} x_{ki} \] (9)

Using the criterion of minimization of the sum-squared error, it is possible to show the estimation of:

\[ \alpha = \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \] (10)

where \( n \) = total number of observation sets (rainfall values).

One needs to estimate \( \beta \) using

\[ \beta_i = S_{xx}^{-1} \cdot S_{xy} \] (11)

where

\[ S_{xy} = \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(y_i - \bar{y}) \] (for \( j = 1, 2, \ldots, k \)) (12)

\[ S_{x_{ji}x_l} = \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)(x_{il} - \bar{x}_l) \] (for \( j = 1, 2, \ldots, k \)

and \( l = 1, 2, \ldots, k \)) (13)

The multiple regression (MR) parameters given in Equation (7) were worked out with the help of identical data used in the training and testing data processed in RBF neural networks. The model so established was used to forecast the same rainfall as in the case of the neural network, where input, which is referred to as the independent variable, is the same as the input to the neural network used, which is the average of each decade as calculated from Equation (4), and the output is the actual rainfall data. A list of MR model parameters \( a \) and \( b \) (only one independent variable in this application) in Equation (7) is reported in the Appendix.

Performance of the rainfall forecasting models

The RBF network model and MR model are designed to forecast the rainfall in Alexandria City. Because there was no definitive test to evaluate the success of each model, a multicriteria assessment was carried out. Basically, the performance of the model is evaluated based on the comparison between the forecasted data and actual data. The prediction of each model is evaluated using the correlation of coefficient \( (R^2) \), root mean square error (RMSE), relative root mean square error (RRMSE), and the absolute percentage error (MAPE). Formulas for calculating \( R^2 \), RMSE, RRMSE and MAPE are given as follows:

\[ R^2 = \frac{\sum_{i=1}^{n} [(D_{a(t)} - D_{f(t)})]/(D_{a(t)} + D_{f(t)})]}{\sum_{i=1}^{n} [(D_{a(t)} - D_{a(t)})^2]} \] (14)

\[ \text{RMSE} = \left[ \frac{1}{n} \sum_{i=1}^{n} (D_{a(t)} - D_{f(t)})^2 \right]^{1/2} \] (15)

\[ \text{RRMSE} = \left[ \frac{1}{n} \sum_{i=1}^{n} \left( \frac{D_{a(t)} - D_{f(t)}}{D_{a(t)}} \right)^2 \right]^{1/2} \] (16)

\[ \text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{D_{a(t)} - D_{f(t)}}{D_{a(t)}} \right) \times 100\% \] (17)

The value of \( R^2 \) of 90% indicates a very satisfactory model performance whereas a value in the range 80–90% indicates a fairly good model (Kachroo 1986). Generally, RMSE and RRMSE formulas evaluated the models based
on a comparison of the estimated error between the actual and forecasted model. A model with the minimum error is considered the best choice. Johnson & King (1988) stated that the MAPE around 30% is considered a reasonable prediction. Further, the analysis will be considered very accurate when the MAPE is in the range of 5–10%.

Finally, the accuracy of the proposed model is measured in terms of coefficient of efficiency (CE), given by Equation (16).

\[
CE = 1 - \frac{\sum_{t=1}^{n}(D_{a(t)} - D_{f(t)})^2}{\sum_{t=1}^{n}(D_{a(t)} - D_{a(t)})^2}
\]

where \(D_{a(t)}\) = the actual rainfall, \(D_{f(t)}\) = the forecast rainfall, \(D_{a(t)}\) = actual mean rainfall, \(D_{f(t)}\) = forecast mean rainfall, \(n\) = number of forecasting periods.

RESULTS AND DISCUSSION

Yearly model predictive rainfall

The topological structure of the RBFN is developed to find the yearly forecasted rainfall using 42 years of rainfall data as training, collected from four rainfall stations in Alexandria City, Egypt, and 8 years as testing data. The neural network is designed by choosing the model construction of the RBF network to reflect the average 10 year rainfall in each decade in order to find the next year. Based on preliminary examination of the data, the RBF network model for yearly forecasted rainfall was built up on an average yearly data base input with actual rainfall data output to train the network in supervised learning techniques where the radius spread used 0.00006 of the RBF function in the network dimension. Figure 9 illustrates the testing error over the 8 years as calculated in the network. It is clear from Figure 9 that the result of yearly forecasting has an error of less than ±10% except one year; overall the maximum error reached –26% in 2006 and for the rest of the years it returned to lower values, less than 5%.

As a comparison for the rainfall forecasting model, a multiple linear regression MR model was used in yearly forecasting. The MR model parameters were calculated on the basis of the network testing data using the MR equations above. In order to examine the rainfall depth in the city, Figure 10 illustrates the observed and forecasted rainfall from the RBF network and MR model. As can be observed from Figure 10, the RBF network models perform best at predicting the rainfall depth, with median difference from the actual rainfall of more than 10 mm, except in 2004 with 52 mm more than actual rainfall. In two of the remaining years the prediction was less than the actual rainfall by 28 and 45 mm in 2003 and 2006, respectively. It can also be seen that, although the MR model provides forecasted rainfall with a difference from the actual rainfall of less than 20 mm, the MR model failed to predict the rainfall in 2004, 2008 and 2009 with a deviation from the actual rainfall of –135, 97 and 42 mm, respectively.

Yearly model efficiency and performance

The values of efficiency, \(R^2\) and CE of each model are presented in Table 1. It is observed that the RBF models outperform the MR model during testing. With a coefficient of efficiency, CE, approaching 1, the RBF model is very highly accurate and reliable for forecasting yearly rainfall.
Comparison of the applied RBF network and the MR model indicated that the RBF performance is superior to the MR, whereas the RMSE and RRMSE of the RBF model are lower than corresponding values of the MR model. The MR gives a higher error than RBF with a poorer degree of efficiency, in term of RMSE and RRMSE. The RMSE for the RBF model is consistently less than half of that for the MR model. The result for the mean absolute percentage error (MAPE), which considers the value of model accuracy, was 13% in the RBF model, which, as it is close to 10%, can be considered as very accurate (Johnson & King 1988). Finally, the values of $R^2$ indicated a good performance of RBF in contrast to the poor value in the MR model.

**Monthly model predictive rainfall**

The January–December forecasts for Alexandria were modeled to forecast the rainfall. The same model construction of the RBF network was used in monthly forecasting which is based on the average 10 year rainfall in each decade to find the next month’s rainfall. The RBF network model for January forecasted rainfall was built up on an average monthly data base input with actual rainfall data output to train the network in supervised learning techniques where the radius SPREAD used 0.00008 of the RBF function in the network dimension. Figure 13 illustrates the testing error over the month of January for 8 years as calculated in the network. It is observed from Figure 13 that the error of the January forecast is between ±20% except for one year, 2002, when overall the maximum error reached 22%; for the remaining years the error returned to lower values of less than 5%. As a comparison for the rainfall forecasting model, a multiple linear regression MR model was used in monthly forecasting. As can be observed from Figure 14, the RBF network model performs better at predicting the rainfall depth, with median difference from the actual rainfall of less than 10 mm, except in 2002 with 23 mm more than actual rainfall. It can be seen that, although the MR model provides forecasted rainfall with a

<table>
<thead>
<tr>
<th>Performance indicators of RBF and MR models</th>
<th>RBF</th>
<th>MR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>27.13</td>
<td>62</td>
</tr>
<tr>
<td>RRMSE</td>
<td>0.15</td>
<td>0.33</td>
</tr>
<tr>
<td>MAPE</td>
<td>13%</td>
<td>21%</td>
</tr>
<tr>
<td>CE</td>
<td>0.972</td>
<td>0.455</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.94</td>
<td>0.2</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.45</td>
<td>1.75</td>
</tr>
<tr>
<td>Mean</td>
<td>207.5</td>
<td>224.9</td>
</tr>
</tbody>
</table>

| Figure 11 | Actual versus forecasted yearly rainfall using RBF model, 2002–2009. |
| Figure 12 | Actual versus forecasted yearly rainfall using MR model, 2002–2009. |
| Figure 13 | Error distribution as a measure for the RBF network for January. |

Table 1: Performance indicators of RBF and MR models
difference from the actual rainfall of less than 30 mm in 2002 and 2004, the rest of the predicted rainfall had a deviation from the actual rainfall of more than 25 mm. The accuracy of the results from both the RBF and the MR models compared with the actual rainfall data is illustrated in Figures 15 and 16. It is clear that the RBF achieved better accuracy \( R^2 = 0.891 \) compared with the MR model \( R^2 = 0.121 \).

The RBF network model for December forecasted rainfall was built up on an average monthly data base input with actual rainfall data output to train the network in supervised learning techniques where the radius \( \text{SPREAD} \) used 0.00004 of the RBF function in the network dimension. Figure 17 illustrates the testing error over the month of December for 8 years as calculated in the network. It is clear from the figure that the error of the December forecast is less than \( \pm 10\% \) except in one year, 2004, when overall the maximum error reached \(-18\%\); for the rest of the years the error returned to lower values of less than 5\%. In order to examine the rainfall depth in the city, Figure 18 illustrates the observed and forecasted rainfall from the RBF network. As shown in Figure 18, the RBF network model performs better at predicting the rainfall depth with median difference from the actual rainfall of less than 5 mm, except in 2004 with 29 mm more than actual rainfall. It can be seen that, although the MR model failed to forecast rainfall, with a difference from the actual rainfall of less than 70 mm in 2004, for the rest of the data the rainfall
predicted had a deviation from the actual rainfall of more than 25 mm.

**Monthly model efficiency and performance**

Comparing the applied RBF network and the MR model (see Table 2) indicated that the RBF performance is superior to the MR for both months, whereas, the values of RMSE and RRMSE for the RBF model are less than those for the MR model. The MR gives a higher error than RBF with a poorer degree of efficiency in terms of RMSE and RRMSE. The RMSE for the RBF model is consistently less than half of the RMSE for the MR model. The result for the mean absolute percentage error (MAPE), which considers the value of model accuracy, was 8.7 and 7.2% in January and December, respectively, in the RBF model, which, at less than 15%, can be considered as very accurate (Johnson & King 1988). Finally, the values of $R^2$ indicated a good performance in RBF compared with the poor value in the MR model.

![Figure 19](https://iwaponline.com/jwcc/article-pdf/3/2/125/375125/125.pdf)

**Figure 19** | Actual versus forecasted December rainfall using RBF, 2002–2009.

Furthermore, it can be observed from Figures 19 and 20 that the RBF outperformed the MR in forecasting the rainfall during the testing session. As depicted in Figure 19, the RBF model achieved $R^2 = 0.986$, while the MR model did not perform so well with $R^2 = 0.338$.

**CONCLUSION**

The purpose of this study is to evaluate the application of neural networks in hydrological forecasting, in particular, rainfall forecasting. The performance of neural networks and their ability to forecast was compared with a statistical model, multiple regression (MR). The RBFN was applied to forecast the rainfall in Alexandria City, Egypt. The primary focus of the research was forecasted yearly rainfall, with additional forecasting of the two highest monsoon rainfall months (January and December) for the same city.

The major conclusions of this research are:

1. The result of the rainfall forecasted based on the RBFN with a single-input single-output model predicted the yearly and monthly values with very high accuracy.
2. The model of multiple regressions did not perform as well as the artificial neural network. In the yearly forecasting, the model gives lower values of $R^2 (0.2)$ and efficiency of error RMSE (62).
3. For yearly forecasting by the RBF network, there was a very good match between the observed values and those predicted values, with $R^2$ value of 0.94 and the value of the efficiency with minimum error RMSE 27.13.
4. The values of rainfall depth for yearly forecasting modeled by RBF give a stable difference from the actual

![Figure 20](https://iwaponline.com/jwcc/article-pdf/3/2/125/375125/125.pdf)

**Figure 20** | Actual versus forecasted December rainfall using MR, 2002–2009.

<table>
<thead>
<tr>
<th>Performance indicators of the RBF and MR models</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>January</strong></td>
</tr>
<tr>
<td>RBF</td>
</tr>
<tr>
<td>RMSE</td>
</tr>
<tr>
<td>RRMSE</td>
</tr>
<tr>
<td>MAPE</td>
</tr>
<tr>
<td>CE</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Mean</td>
</tr>
</tbody>
</table>
rainfall with a minimum of 0.7 mm, except one year with a maximum difference near 50 mm.

5. The RBF network has been proven to be a robust model in forecasting the yearly rainfall with a coefficient of efficiency CE equal to 0.97. However, the MR model did not forecast as well as RBF with a lower value of CE of 0.45.

6. Monthly forecasting based on RBF networks were as accurate as the yearly model analyzed. For January the forecasted model $R^2$ is 0.89 and the value of the efficiency of error RMSE equal to 10.61. However, the result of the December forecast is more accurate than January with values of $R^2$ and RMSE of 0.98 and 10.85, respectively.

7. For January rainfall depth started with a minimum difference equal to 2.8 mm reaching a maximum in one year of 23 mm. December rainfall depth had a minimum difference equal to 0.7 mm for all years except one where the difference reached 29 mm.

8. The RBF network gives a high efficiency for both months compared with the MR model, with a coefficient of efficiency (CE) for January equal to 0.94 and for December, 0.99. The CE is 0.35 and 0.58 for January and December, respectively, in the MR model forecast.

Based on the results of this research, the RBF neural network model has significantly better stability than the MR model. Despite the fact that using artificial intelligent modeling in forecasting was a reliable way to make a conservative estimate for the actual monitoring data, overall there may be a limit on the RBF neural network. The major limitation is the trial and error procedure for selecting the optimal structure of the network. Future research should integrate the neural network model with an advanced optimization model such as a genetic algorithm in order to search for the optimal neural network structure. Temperature and wind speed data could be used as input data to evaluate the model performance in forecasting.

ACKNOWLEDGEMENTS

This research is supported by: a research grant to the first author by Smart Engineering System, University Kebangsaan Malaysia; and eScience Fund project 01-01-02-SF0581, Ministry of Science, Technology and Innovation, (MOSTI) and UKM-DLP-2011-02.

REFERENCES


First received 30 January 2011; accepted in revised form 16 September 2011. Available online April 2012

**APPENDIX: MULTI-REGRESSION MODEL**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>a</th>
<th>b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model period</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yearly</td>
<td>164.15</td>
<td>0.2072</td>
</tr>
<tr>
<td>January</td>
<td>56.105</td>
<td>0.1249</td>
</tr>
<tr>
<td>February</td>
<td>40.959</td>
<td>0.338</td>
</tr>
</tbody>
</table>