Journal of Engineering and Sustainable Development
www.jeasd.org
Vol. 20, No.06, November 2016
ISSN 2520-0917

MONTHLY RAINFALL QUANTITIES FORCASTING USING NARX NETWORK

*Dr. Mohammed Ali Tawfeeq[1], Ghusoon Idan Arb[2]

1) Assist Prof., Comp. and Software Eng. Department, Al-Mustansiriayah University, Baghdad, Iraq.
2) Assist Lecturer, Comp. and Software Eng. Department, Al-Mustansiriayah University, Baghdad, Iraq.

Abstract: An accurate precipitation forecast can reflect positive impact in several areas. It provides helpful data in hydrological projects designs, such as constructing dams, reservoirs, rainfall networks, as well as takes some precautionary measures that can overcome the flooding problems. This paper proposes a monthly quantitative precipitation forecasting model that covers the total land area of Iraq. The model is based on the use of Nonlinear AutoRegressive with eXogenous input neural network (NARX). This type of network is considered as one of the most important dynamic networks that can deal with time series data. It is a type of recurrent networks with feedback connections between its layers and a tapped delay lines. The data used to train and test the network are real data obtained by NASA GES DISC which represent monthly quantitative precipitation of more than 1350 site uniformly distributed to cover the land of Iraq for a historical period of ten years. The designed forecasting network model showed good performance, in which the total calculated MSE for the testing data set is about (2.8×10^{-3}), and the its correlation coefficient R is about (0.95). The correlation of the predicted error with time has been checked also; it showed that almost all the autocorrelation function values are fall within the bound of the confidence interval.

Keywords: *Forecasting, Precipitation, NARX, Recurrent neural networks*

تنبؤ كميات الامطار الهاطلة شهرياً بأستخدام شبكات التغذية العكسية الديناميكية العصبية

الخلاصة: ان التنبؤ الدقيق لكميات هطول الأمطار يمكن ان ينعكس إيجابا وبشكل مؤثر في العديد من المجالات. حيث أنه يوفر بيانات مساعدة عند اعداد تصاميم المشاريع الهيدرولوجية، مثل بناء السدود والخزانات وشبكات مياه الأمطار، وكذلك لاتخاذ بعض التدابير الاحترازية التي يمكن من خلالها التغلب على مشاكل الفيضانات. اقترح هذا البحث نموذج تنبؤ كمي شهري لهطول الامطار وبما يغطي اجمالي مساحة العراق. يستند هذا النموذج على استخدام شبكات عصبية ذات انحدار غير خطي مع مدخلات خارجية المنشأ (NARX). ان نوع هذه الشبكة يعتبر أحد الشبكات الديناميكية الأكثر أهمية التي يمكن أن تتعامل مع بيانات السلاسل الزمنية. ان هذا النوع يمثل الشبكات ذات النواتج المرتدة عكسياً والتي تحوي على وصلات تغذية عكسية بين طبقاتها مع خطوط تأخير توظيف نواتجها. البيانات المستخدمة لتدريب واختبار الشبكة هي بيانات حقيقية تم الحصول عليها من موقع وكالة ناسا غيس دسك (NASA GES DISC) والتي تمثل كميات الامطار الهاطلة شهرياً لأكثر من ١٣٥٠ موقع موزعة بشكل متجانس لتغطية مساحة العراق لفترة تاريخية تمتد لعشر سنوات. أظهر نموذج الشبكة التنبؤي المصمم أداء الجيداً، وقد بلغت كمية معدل مربع الخطأ باستخدام بيانات الفحص حوالي ($2.8*١٠^{-٣}$)، واجمالي معامل الارتباط R هو حوالي (0.95). وقد تم التحقق ان كان هنالك ترابط ما بين الخطأ المتوقع والزمن: حيث أظهرت نتائج احتساب الارتباط الذاتي ان جميع القيم تقريبا تقع ضمن حدود فترة الثقة.

* drmatawfiq@yahoo.com

1. Introduction

The prior knowledge and the good prediction can significantly help in the development of scientific plans that can contribute to the creation of suitable infrastructure to serve communities.

Forecasting of precipitation introduces valuable information to the planners and engineers to put forward appropriate strategies for hydrological solutions, such as the construction of dams, reservoirs and basins, drainage and pumping stations, etc.

The importance of this issue is the main motivation for several researchers to introduce their approaches about Quantitative Precipitation Forecasting (QPF). In general these approaches are based on probabilities, statistical analysis, extrapolation, clustering analysis, neural networks, and others. For the importance posed by Artificial Neural Networks (ANN) and their potentials in the field of predicting future values, most of the available up-to-date approaches dealt with the subject of QPF are based on the use of NN techniques.

K. CROWELL, [1] introduces two different ANN models to predict both the probability of precipitation and quantitative precipitation over a 24-hour period, the first model uses genetic algorithm to calibrate the data for selection, while the selected data for the second one are not calibrated. The ANN models use a data pattern classification algorithm known as a Probabilistic Neural Network (PNN) with one hidden layer and output layer. The author claims found that both models show unacceptable overall accuracy with the evaluation data set. N. Q. Hung et. al. [2] compare and evaluate six Multilayer Perceptron (MLP) neural network models with continuous data for rainy and non-rainy days. The networks are trained using backpropagation with momentum algorithm and training data of the waterfall for 3 consecutive years (1997 – 1999), with meteorological parameters for 75 rain gauge station in Bangkok, Thailand. Among the six NN models, the generalized feedforward network with two hidden layers and hyperbolic tangent activation function is assumed to be most satisfactory in forecasting precipitation. This research concludes that the most important data beside the rainfall is the wet bulb temperature.

H.D.P. Weerasinghe et. al., [3] consider a neural network model to cluster ten weather stations in the dry zone of Sri Lanka. The authors test several feed-forward neural network architectures with backpropagation learning algorithm. They cliam that the 12-11-1 network architecture show the best performance, with accuracy of 79±3% for predicting the daily precipitation. N. Khalili et. al.,[4] propose a three-layer feedforward neural network trained by gradient decent backpropagation algorithm to forecast daily rainfall in Mashhad Synoptic Station. They use the data of March, May, and December for the period from 1986 to 2010. Kostas P. Moustris et. al., [5] forecast the monthly maximum, minimum, mean and cumulative precipitation totals within a period of the next four consecutive months. The precipitation datasets concern monthly totals recorded at four meteorological stations in Greece for a 115-year period (1891–2005). The authors use 16 ANNs, four for each station. The structure of each net is 7-5-1 with backpropagation training algorithm. They point that the developed ANNs did not have the ability to forecast the peaks in all cases. T. Santhanam and A.C. Subhajini [6] use Radial Basis Function (RBF) NN and backpropagation NN for weather forecasting. They point that the performance of the RBF network is better than the backpropagation netwok; they remark the prediction accuracy for the RBF is 88.49%. G. Shrivastava et.al. [7] presented a survey on the use of ANNs in weather forecasting. They conclude that the backpropagation networks and radial basis function are efficient enough to forecast monsoon rainfall, and other weather parameters over small geographical region. Another survey introduced by D.R. Nayak et. al. [8]. They reported that, the use of ANN techniques in rainfall prediction is more suitable than traditional statistical and numerical methods**.** F. Mekanik et. al. [9] model rainfall using multiple regression and ANN learned by Levenberg–Marquardt algorithm for Victoria, Australia. They predict in advance three years spring rainfall. They cliam that NN models are more accurate than multiple regressions for long term prediction. p. Gupta et. al. [10] develop different ANN topologies for rainfall time series forecasting. In their evaluation to their different models, they found the backpropagation training algorithm is more accurate than other traditional models. Taksande et al., [11] used ANN with backpropagation learning and genetic algorithm (GA) for rainfall. In GA, they use the Hidden Markov Model (HMM) to record the previous data. Based on their experiment result, using several population size and crossover probabilities, they found that the combination of GA and HMM can gives prediction graph with higher than 90% accuracy.

The behavior of the precipitation is completely non-linear and not easy to predict. This arise the need to use a sophisticated methodology that can introduce accurate forecasting. This work suggests solving this problem by using a Nonlinear Auto Regressive recurrent with exogenous input neural network (NARX). Although this type of ANN is complicated, but it can model and predict non-linear time series data efficiently.

The rest of this paper is organized as follows; Section 2 introduces an idea about dynamic recurrent networks, Section 3 describes the proposed recurrent neural network used for forecasting precipitation. It also illustrates the training data and the preprocessing operations. The experimental results are introduced in section 4. Section 5 contains the conclusion of this work.

2. Dynamic Recurrent Networks

Neural networks can be classified as either static networks or dynamic networks. The traditional feedforward neural networks such as MLP, RBF, Self Organizing Map (SOM), etc. are static type. The input data to the static networks is of type concurrent. On the other hand, dynamic networks operate on a sequence of input data, i. e. on a data of type sequential that is varied with time. The architecture of the dynamic networks can be either of feedforward connections with time delayed units at the input connections, or it can be a recurrent with feedback connections, or of both of them. The time delay units and the feedback connections cause the dynamic network to have memory. The output of such type of networks at any given time depends on the current input and the previous values of the input sequence. Recurrent networks may have different architectures and different models. S. Haykin [12] describes four main network models: input-output recurrent model, state space model, recurrent multilayer perceptron, and second order networks. Each of these models represents a specific form of global feedback network. Hagan et. al. [13] introduces Layered Digital Dynamic Networks (LDDN). It is a general representation of dynamic network that can be used to express multiple recurrent connections with tapped delay lines networks.

Dynamic recurrent networks can be implemented effectively using nonlinear time series data in prediction problems. Prediction using such networks can be classified to three main categories:

1- Nonlinear input output: Predict series $y(n+1)$ given present and p past values of input series $u(n), \dots, u(n-p+1)$. The behavior of such network can be represented by

$$y(n+1) = F(u(n), u(n-1), \dots, u(n-p+1)) \quad (1)$$

2- Nonlinear autoregressive (NAR): The prediction of series $y(n+1)$ depends on current and p past values of $y(n)$. this system can be described by

$$y(n+1) = F(y(n), y(n-1), \dots, y(n-p+1)) \quad (2)$$

3- Nonlinear autoregressive with exogenous inputs (NARX): Predict future series $y(n+1)$ based on two series:
   - Input series u composed of present and p past values, $u(n), \dots, u(n-p_u+1)$.
   - Delayed output series values, $y(n), \dots, y(n-p_y+1)$.

   The appropriate dynamic network to be used for solving such problems is the NARX network shown in Figure 1, in which it consists of two tapped delay lines (TDL), one is at the input side and the other is at the feedback connections. The behavior of NARX can be described by [12]

$$y(n+1) = F(y(n), \dots, y(n-p_y+1); u(n), \dots, u(n-p_u+1)) \quad (3)$$

Where F in the above three equations is a nonlinear function.

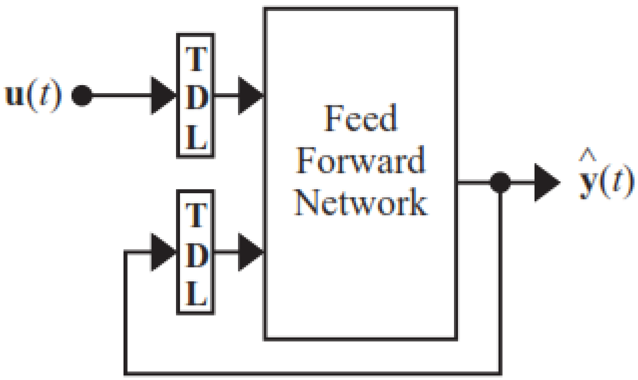

Figure-1 Nonlinear autoregressive with exogenous inputs [14]

There are two training modes for recurrent networks; epochwise and continuous training. In epochwise mode, temporal sequences of input/target response pairs are selected to be used in the training. The training starts by considering some initial state, and stops when the network reaches a new state to start again with new initial state for the next epoch. On the other hand, the continuous training mode is more suitable for on-line learning, in which the networks learn during its operation and the learning never stops [12].

Among the different training algorithms, the standard backpropagation learning process can be modified to accommodate the requirement of adjusting the weights and biases of dynamic networks. There are two different approaches to do this; BPTT and Real-Time Recurrent Learning (RTRL). BPTT algorithm proposes unfolded multilayer feedforward network for some temporal operation of the dynamic network. Then the standard backpropagation training algorithm is applied to adjust the network parameters by determining the local error gradient at the output using the partial derivative of the performance index, then propagate this error backward with time. BPTT can be implemented using epochwise mode, continuous mode, or in combination of the two modes.

For epochwise BPTT, the cost function E_{Total} may be define by [12]

$$E_{Total} = \frac{1}{2} \sum_{n=n_0}^{n1} \sum_{j \in N} e_{j,n}^{2} \tag{4}$$

Where $e^{2}_{j,n}$ is the squared error at the output of neuron j at time n, N is the number of the neurons in the network, and the epoch starting and ending time are n_0, n_1 respectively.

The local gradient error at the output is the partial derivative of the cost function with respect to local induced field v of neuron j at time n is given by

$$\delta_{j.n} = -\frac{\partial E_{Total}}{\partial v_{j,n}} \tag{5}$$

for all $j \in N$ and $n_0 < n \leq n_1$. The updating of synaptic weight w_{ji} of neuron j is

$$\Delta w_{ji} = -\alpha \frac{\partial E_{Total}}{\partial w_{ji}}$$

$$= \alpha \sum_{n=n_0+1}^{n_1} \delta_{j,n} x_{i,n-1} \tag{6}$$

where, α is the learning rate parameter and $x_{i,n-1}$ is the input applied to neuron j.

In RTRL algorithm the derivatives are calculated at each time step and propagated forwarded through time, thus, it is well suited for real time network applications [13].

3. Proposed Precipitation Forecasting Recurrent Network

In general, the field of forecasting needs to deal with time series data. Precipitation forecasting requires a historical time series data which is completely nonlinear. Accordingly, this requires a complex methodology for accurate forecasting. In this work, a recurrent neural network is proposed to solve this problem. More specifically a Nonlinear Auto Regressive with exogenous inputs NARX network is proposed. The NARX network model uses past inputs and past outputs to forecast the precipitation of the future values.

This paper assumes a monthly quantitative precipitation forecasting model for the total land of Iraq which is spanning more than 437,000 km2 with varied topography. In such circumstances, and for accurate forecasting, the neural network should know the previous behavior of the precipitation for several years. This work considers the availability of data for previous year to predict the amount of precipitation for the next year.

The proposed architecture of the network is composed of input layer, one hidden layer, and output layer with feedback connection as shown in Figure 2. Each input vector presented to the network consists of four elements; latitude, longitude, year, and the accumulated rainfall. Thus the input layer has four input units. The best number of hidden neurons empirically found is 24 neurons, and the output layer has just one neuron which produces the forecasted future precipitation value. The tapped delay lines (TDL) at the input side and at the feedback connections are assumed to be of length two, so as to give the network the ability to identify suitable future values during the training stage, and at the same time not to make the network of high complexity.

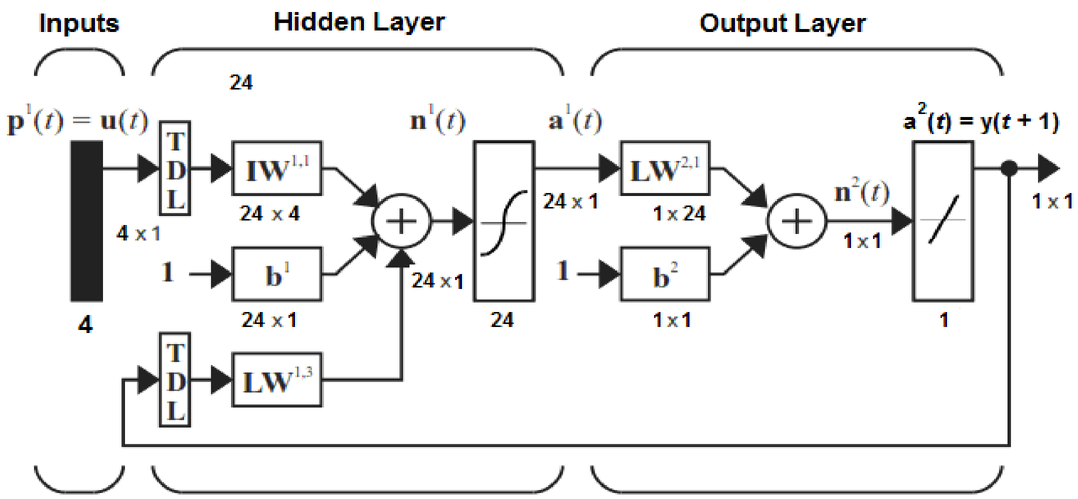

Figure-2 Proposed NARX network architecture

The activation function used with the hidden layer neurons is the hyperbolic tangent function, and for the output neuron is the identity function.

The network is trained using Backpropagtion Though Time (BPTT) in epochwise mode. The training utilizes Levenberg-Marquardt backpropagation algorithm.

Comparing the proposed NARX network with Elman’s recurrent network, Elman’s network is assumed as a simple recurrent network. It stores the output of the hidden neurons for one time-step only to feed them back to the input layer [12], so that it has only one TDL.

3.1. Data Preprocessing

One of the main issues in neural networks is the data selected to train the network. Generally, the training data should span the input/target space adequately. In this work it is required to train the proposed NN to cover the land of Iraq. The training data have been gathered from NASA GES DISC [15]. These data represent the monthly rainfall quantities for a total of 1353 site lies between latitudes 29' and 37' north and between longitudes 38' and 48' east, for the time period from 2004 to 2013. A sample of this data is shown in Table 1. The data set for the first eight years is used to train the proposed network, while the data set for the last two years is used to test the trained network.

The gathered data are classified and grouped according to the month they represent. As a result, a twelve data set are prepared, each set represent a time series of the rainfall values for the selected time period. The reason behind this data separation is due to vast disparity in rainfall quantities per each month, as well as, the big difference in the nature of Iraq topography. Each data set is used to train a NARX network which has been used later for precipitation forecasting of the month represented by the training data.

Table 1 Sample of the used data

Latitude	*Longitude*	*Accumulated Rain*	*Month/ Year*
29	38	2.6294	April/2004
30	38	0	July/2005
30	44	0.2652	August/2006
32	48	67.4601	December/2007
33	38	4.7609	December/2008
34	48	4.62	June/2009
34	48	29.4989	May/2010
36	46	44.958	November/2011
36	48	11.2635	October/2012
37	38	2.7445	September/2013

For good training results, each data set is partitioned to three subsets; the first subset is for training and the other two subsets are for validation and test. According to the work in [13], the training subset has been selected randomly to approximately 70% of the total training data set, while the validation and test subsets are 15% each.

In order to facilitate the training process and to reduce the output error, it is preferable to normalize the training data. Thus, before starting the training process the data have been normalized to be within zero mean and variance of 1. This is done by

$$d_n = \frac{d - \bar{d}}{\sigma_d} \tag{7}$$

Where d_n is the normalized value of d, and $\bar{d}$ is the mean of the data, while σ_d is its standard deviation.

3.2. Experimental Results

After the training is finished, the network need to be analyzed and tested to check its performance. The main analysis and test techniques applied to the trained network, taking December dataset series from 2004-2013 as example, are the following:

1- The obtained Mean Square Error (MSE) between the actual response and the target for the testing data is 2.8×10^{-3}. In general, the desired MSE is assumed to be zero which is practically unreachable in such nonlinear input/ output forecasting, hence the obtained error can be considered negligible comparing with the zero MSE.

2- The correlation coefficient R between the actual outputs and the target values for the testing data is computed. For ideal performance, when there were no error at the output the value of this coefficient should be 1. Here, the computed value of R using the testing data is about 0.95.

3- Network training performance plot which shows the training progress and plot the MSE at each epoch is illustrated in Figure 3. In which, although the training error keeps going down but the training stop at the point where the validation error record the minimum error. This is to prevent over fitting in the training process. The degree of the network generalization can be determined according to test subset error which is found here a relative small value.

4- Regressions between the actual output and the target values for the test subset data is plotted as shown in Figure-4.

5- Test the correlation of the prediction errors in time using the autocorrelation function. The test result is illustrated in Figure-5. The result showed values close to zero and fall within the bounds of the confidence intervals except for lag = 0 and the starting and ending lags, which means that the prediction errors are not correlated and the network is trained successfully.

6- The error verses time for each output/ target pair for the overall data is plotted as shown in Figure-6. Although this plot is crowded even when zooming in, but it can be noticed that the outputs are very close to the target values. This can be presented more clearly by plotting the error histogram as shown in Figure-7

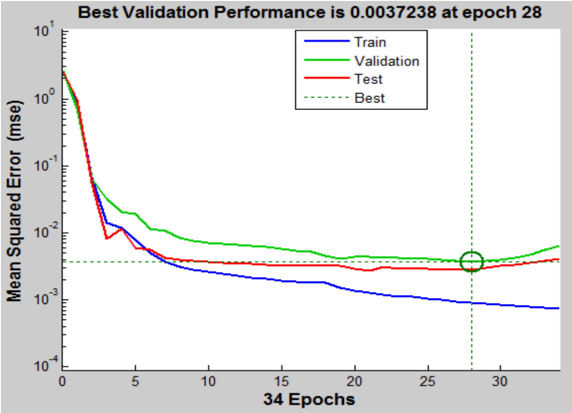

Figure-3 Neural network training performance

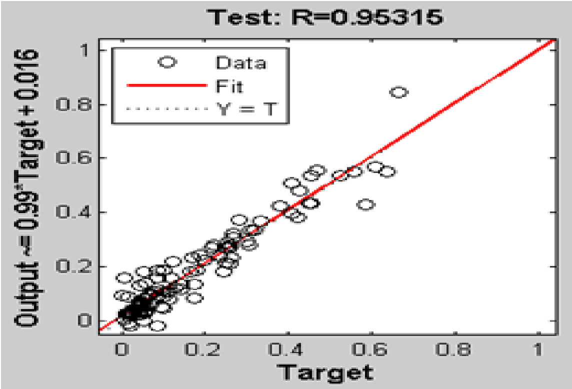

Figure-4 Regression between actual outputs and targets

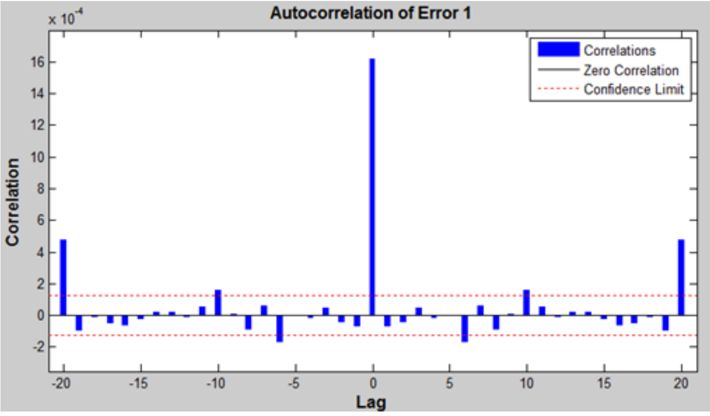

Figure-5 Autocorrelation of prediction errors in time

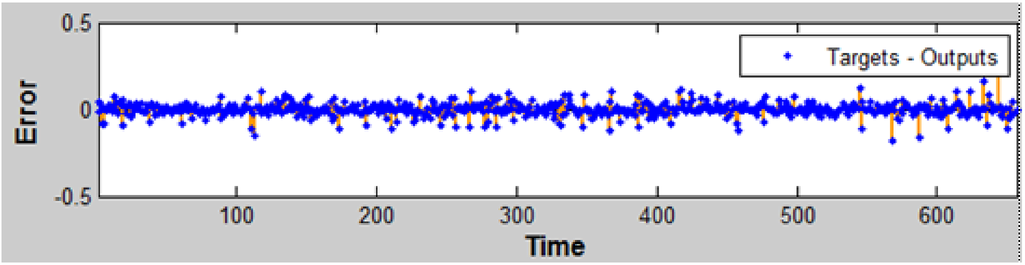

Figure-6 Time series response

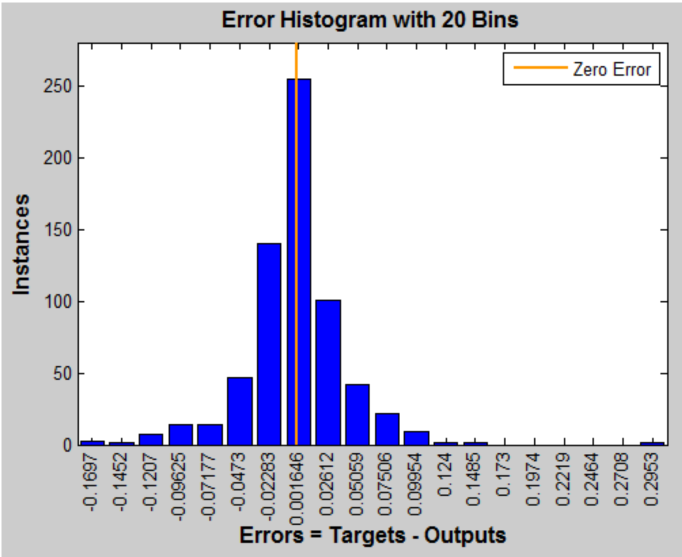

Figure-7 Error histogram

4. Conclusions

In general the dynamic recurrent NNs can deal with time series data. In this work a recurrent network of type NARX is used for precipitation forecasting that covers the total land of Iraq. The time series data used represent the monthly quantitative precipitation in Iraq for the period from 2004–2013. These data is divided into 12 groups; each group represents the quantity of the rainfall for particular month. Accordingly, a standalone network of the same design is trained to forecast the precipitation of each month. The designed model consists of four units input layer with TDL of length two, hidden layer composed of 24 neurons, and output layer of single neuron and TDL of length two. The training algorithm used to train the net is BPTT in epochwise mode.

The performance of the designed network model has been analyzed and tested. It is found that the overall MSE for the testing data (using the dataset of December as example) is about 2.8×10^{-3} with correlation coefficient of more than 0.95. The degree of the network generalization is ensured, this is due to using the validation technique in the training process to prevent data overfitting. The regression between the actual responses and the target values is found closely related to each other. On the other hand, it is found that the error is not correlated with time, which means that the network is trained successfully.

5. References

1. Kevin L. Crowell (2008). "Precipitation Prediction Using Artificial Neural Networks", M.Sc. Thesis, ATHENS, GEORGIA.
2. N. Q. Hung, M. S. Babel, S. Weesakul, N. K. Tripathi (2009). "An Artificial Neural Network Model For Rainfall Forecasting In Bangkok, Thailand", Hydrology and Earth System Sciences, 13, , pp. 1413–1425.
3. H.D.P. Weerasinghe, H.L. Premaratne, and D.U.J. Sonnadara (2010). "Performance of neural networks in forecasting daily precipitation using multiple sources", .Natn.Sci.Foundation Sri Lanka 38 (3), pp. 163-170.
4. N. Khalili, Khodashenas Saeed Reza, Davary Kamran and Karimaldini Fatemeh (2011). “Daily Rainfall Forecasting for Mashhad Synoptic Station using Artificial Neural Networks”, International Conference on Environmental and Computer Science, vol. 19, pp. 118-123.
5. Kostas P. Moustris, Ioanna K. Larissi, Panagiotis T. Nastos, and Athanasios G. Paliatsos (2011). "Precipitation Forecast Using Artificial Neural Networks in Specific Regions of Greece", WATER RESOURCES MANAGEMENT.
6. Tiruvenkadam Santhanam and A.C. Subhajini (2011). "An Efficient Weather Forecasting System using Radial Basis Function Neural Network", Journal of Computer Science 7 (7): 962-966.
7. Gyanesh Shrivastava, Sanjeev Karmakar, and Manoj Kumar Kowar (2012). "Application of Artificial Neural Networks in Weather Forecasting: A Comprehensive Literature Review", International Journal of Computer Applications (0975 – 8887), Volume 51– No.18.
8. Deepak Ranjan Nayak, Amitav Mahapatra, and Pranati Mishra (2013). " A Survey on Rainfall Prediction using Artificial Neural Network", International Journal of Computer Applications (0975 – 8887), Volume 72– No.16.
9. F. Mekanik , M.A. Imteaz, S. Gato-Trinidad, and Elmahdi (2013). " Multiple regression and Artificial Neural Network for long-term rainfall forecasting using large scale climate modes", Journal of Hydrology, Volume 503, 30 October.
10. Prince Gupta, Satanand Mishra, and S.K.Pandey (2014). " Time Series Data Mining in Rainfall Forecasting Using Artificial Neural Network", International Journal of Scientific Engineering and Technology, Volume No.3 Issue No.8, Aug.
11. Amruta A.Taksande, Dr. S.P.Khandait, and Prof. Manish Katkar (2014). "Rainfall Forecasting Using Artificial Neural Network: A Data Mining Approach", International Journal Of Engineering Sciences & Research Technology, 3(4); April.
12. Simon Haykin (2009). "Neural Networks and Learning Machines", Pearson International Edition, 3rd Edition.

13. Martin T. Hagan, Howard B. Demuth, M.H. Beale, and Orlando De Jesus (2014). "Neural Network Design", Copyright by Martin T. Hagan and Howard B. Demuth, 2nd Edition.
14. Martin Hagan, Howard Demuth, and Mark Beale (2009). "Neural Network Toolbox™ User's Guide", © COPYRIGHT 1992–2009 by The MathWorks, Inc.
15. National Aeronautics and Space Administration NASA, Goddard Earth Science Data & Information Service Center GES DISC. http://disc.sci.gsfc.nasa.gov/precipitation.