The International Journal Of Engineering And Science (IJES)
||Volume||2 ||Issue|| 6 ||Pages|| 63-73||2013||
ISSN (e): 2319 – 1813 ISSN (p): 2319 – 1805

Short-Term Forecasting Of Dadin-Kowa Reservoir Inflow Using Artificial Neural Network.

[1]Ibeje, A.O. And [2]Okoro, B.C.
[1]Department of Civil Engineering, Imo State University, Owerri, Nigeria
[2]Department of Civil Engineering, Federal University of Technology, Owerri, Nigeria

ABSTRACT

In the study, Artificial Neural Network (ANN) model was applied to forecast the daily inflow into Dadin-Kowa Reservoir along River Gongola in Northern Nigeria. An effective algorithm for daily reservoir inflow predictions, which solicited the observed precipitation, forecasted precipitation from Quantitative Precipitation Forecast (QPF) as predictors and discharges as predicted targets for Multilayer Perception Artificial Neural Networks (MLP-ANNs) modelling, was presented. With a learning rate of ***0.01*** *and momentum coefficient of* ***0.85****, the MLP-ANN model was developed using* ***1*** *input node,* ***7*** *hidden nodes,* ***1000*** *training epoches and* ***24*** *adjustable parameters. Error measures such as the Coefficient of Efficiency (CE), the Mean Absolute Error (MAE), the Mean Squared Relative Error (MSRE), the Relative Volume Error (RVE) and the coefficient of determination (R^2) were employed to evaluate the performance of the proposed model for the month of August. The result revealed: CE= 0.991; MAE= $9.679 X 10^{-6}$; MSRE = $2.49 X 10^{-7}$; RVE = $3.840 X 10^{-6}$ and $R^2 = (0.991)^2$, which showed that the proposed model was capable of obtaining satisfactory forecasting not only in goodness of fit but also in generalization.*

KEY WORDS: *Artificial Neural Network; Proposed Model; Reservoir Inflow; Forecasting*

Date of Submission: 22 May 2013, Date of Publication: 25.June.2013

I. INTRODUCTION

Dadin-Kowa dam was commissioned in 1988 for irrigation, domestic water supply and flood control (Ibeje et al., 2012). Over the years, it has become very difficult to determine the area of cultivable land in each year because there is no prior information of available reservoir inflow. Sometimes the water needs of the cultivable area would be more than the available water in the reservoir. This has often resulted in the reduction of the cultivable area which in turn reduced the amount of agricultural produce. At other times, especially in wet years, the cultivable area would be limited. This resulted in evacuation of some water from the reservoir through the dam outlets. It is therefore very important to forecast the reservoir inflow in order to determine the optimal cultivable area which the reservoir supplies water. Dadin-kowa reservoir has lost large amount of water many times in the recent years. Excess rainfall during rainy season can fill the reservoir and make it to overflow at the end of rainy season. By forecasting the reservoir inflow, the excess water in rainy season could be used to generate hydropower energy before overflowing the dam. The objective of the study is to develop rainfall-inflow model using Artificial Neural Network (ANN).

Artificial Neural Networks

An ANN is a mathematical model which has a highly connected structure similar the brain cells. They consist of a number of neurons arranged in different layers an input layer; on out layer and one or more hidden layers. The input neurons receive and process the input signals and sends an output signal to other neurons in the network. Each neuron can be connected to the other neurons and has an activation function and a threshold function, which can be continuous, linear or non-linear functions. The signal passing through a neuron is transformed by weights which modify the functions and thus the output signal reaches the following neuron. Modifying the weights for all neurons in the network, changes the output. Once the architecture of the network is defined, weights are calculated so as to represent the desired output through a learning process where the ANN is trained to obtain the expected results. Information available is used to define learning or training data set and a validation data set (Rumelhert et al., 1986). Several different architectures and topologies have been proposed for ANN. They differ in terms of architecture, in the learning process and in the training strategy (Nussbaum et al., 1996). A linear model can be represented adequately by a single layer network, while a nonlinear model is generally associated with a multiple layer network. (LS WC, 1999).The use of ANN techniques in water resources and stream flow prediction is relatively new and has been reported by French et al., (1992); Zurada (1992); Hall and Minns (1993); Zealand et al.,(1999); Abrahart et al., (1998); Zhu and Fugita (1994); Hsu et al (1993); Abrahart and See (1998); Minns (1998) and Salazar et al., (1998), among others. Artificial neural Networks have a structure where nonlinear function are present and the parameter identification are based on techniques which search for global maximum in the space of feasible parameter values, and hence can represent the nonlinear effects present in the rainfall-runoff processes. An important advantage of ANN compared to classical stochastic models are that they do not require variables to be stationary and normally distributed (Burke, 1991). Non-stationary effects present in global phenomena, in morphological changes in rivers and other can be captured by the inner structure of ANN (Dandy and Mainer, 1996). Furthermore, ANNs are relatively stable with respect to noise in the data and have a food generalization potential to represent input-output relationships (Zealand et al., 1999).

The Study Area

Dadin-kowa town is located between latitudes 10 to 10^o 20^o N and longitudes 11^o01^oE and 11^o19^oE (Figure 1) it shares common boundary with Akko L.G.A in both the south and west, Yamatu-Deda to the East and Kwami to the North. The climate of Dadin-kowa is characterized by a dry season of six months, alternating with a six months rainy season. As in other part of Nigerian savanna, the precipitation distribution is mainly triggered by a seasonal shift of the inter-tropical Convergence Zone (ITCZ). For the years 1977 to 1995, the mean annual precipitation is 835mm and the mean annual temperature is about 26^oC whereas relative humidity has same pattern being 94% in August and dropping to less than 10% during the harmattan period (Dadin-kowa L.G.A., 1999). The relief of the town ranges between 650m in the western part to 370m in the eastern parts. Dadin-Kowa Dam is a multipurpose dam which impounds a large reservoir of water from Gongola River. It has a storage capacity of 1.77 billion cubic meters for irrigation to $950km^2$ (Ibeje, et al, 2012). Its flood spillway has a capacity of $1.111m^3/s$.

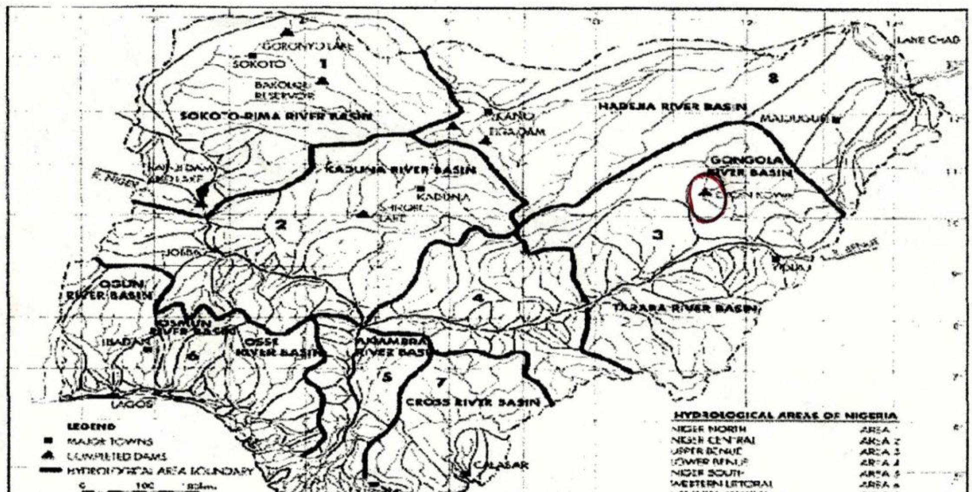


Figure 1: Hydrological Map of Nigeria Showing the Location of Dandi Kowa Reservoir

The study area is the Dadin-Kowa reservoir. This is located at the narrow section of the Gongola River in the present Gombe state, Nigeria (see Figure 1). The dam is a multipurpose project designed to serve among other uses, irrigation, industrial and domestic supply and flood control. Downstream of the River is located a rice farm that is irrigated by a canal from the dam (Ibeje et al., 2012).

Dadin-Kowa Reservoir Inflow Characteristics

Figures 2 and 3 show the variation in the inflow to Dadin-kowa Reservoir. The inflow increased gradually after the commissioning of the dam in 1988 until there was a decline in the inter-annual inflow in 1997. The effect of this is the inability to fill the reservoir due to siltation. The flow also exhibited a noisy pattern. The annual inflow showed that the reservoir inflow increased from the month of May to August, after which it declined to the month of September. The month of August is notably the wettest month. Other months in the year never had any reasonable records of inflow over the years. This is a clear demonstration of the climatic characteristics of the reservoir catchment area. There were usually no rainfalls during those months of no inflows.

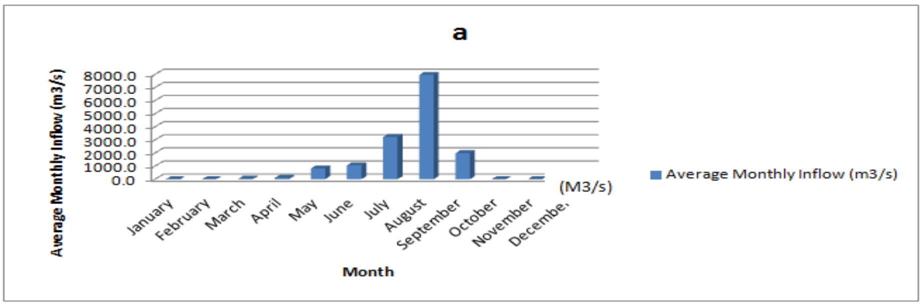


Figure 2: Annual Inflow of Dadin-Kowa Reservoir

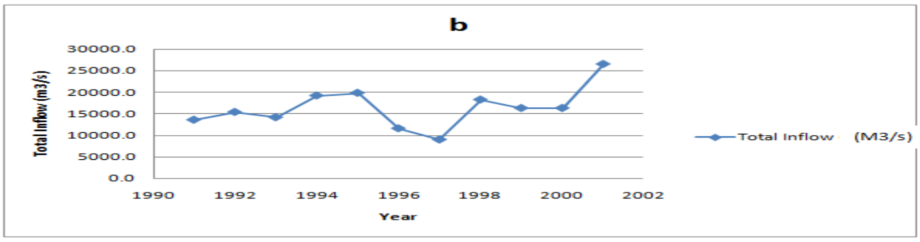


Figure 3: Seasonal Inflow of Dadin-Kowa Reservoir

II. METHODOLOGY

Data scaling is a very important step before the models can be formulated in this study. All the input variables were standardized by subtracting mean and dividing the difference by the standard deviation. This would generate a set of standard normal random variables with mean '0' and standard deviation '1'. The standardized data is sometimes referred as Z – score and is calculated according to the equation.

$$z_x = \frac{x_i - X}{\sigma_x} \qquad (1)$$

$$\dot{X} = \frac{\sum_i^n x_i}{n} \qquad (2)$$

$$\delta x = \sqrt{\frac{\sum_i^n (x_i - x)^2}{n-1}} \qquad (3)$$

Where and X_i = data point, $\dot{X}$ = mean of data set, n = number of data points, δ_x = standard deviation

There are three variables which are relevant in the model. They include: daily inflow, namely: the antecedent observation of rainfall and discharge and the seasonal information. For every day t, four MLP – ANN models, namely: Model (t + i) i = 0, 1, 2 3 are developed for the daily reservoir inflow forecasting with lead – times (represented by symbol i) varying from 1 to 6 days. The model (t+i) i = 0, 1, 2 makes use of the different QPF for the first three days whereas the Model (t+3) is a Unified model for the next four days since no QPF more than three days is available at present.

So, the models can be demonstrated by four equations as follows:

$Model\ (t+0)$:

$$Q(t+0) = f(P(t-2), P(t-1), QPF(t), P(30), Q(30)) \qquad (4)$$

$Model\ (t+1)$:

$$Q(t+1) = f(P(t-2), P(t-1), QPF(t), QPF(t+1)\ P(30), Q(30)) \qquad (5)$$

$Model\ (t+2)$:

$$Q(t+2) = f(P(t-2), P(t-1), QPF(t), QPF(t+1), QPF(t+2), P(30), Q(30)) \qquad (6)$$

$Model\ (t+3)$:

$$Q(t+i) = Model\ (t+0) i = 3,4,5,6; QPF\ (t) = 0 \qquad (7)$$

In which QPF is the quantitative precipitation forecasting P and Q are the antecedent observed rainfall and discharge, P(n) and Q(n) are the mean values of the observed rainfall and discharge of antecedent of days as the seasonal information.

Preparation of inputs for calibrations and verification of models requires splitting the datasets into two sub-datasets in model development. The first sub-dataset contained about two-third of the original data, which was used for the calibration of the inflow forecasting models. The second sub-dataset, which is unused (unseen) during the model calibrations, was thus prepared solely for the verification or validation of the inflow forecasting models. Within such a two-stage effort of model development, models have to be created based on the calibration of datasets, and then verified by the verification dataset. The ultimate predication accuracy may be confirmed in the verification. A statistic approach, proposed by Sudhear (2002), is used here to select the antecedent predictors which have the most significant effect on the inflow of following days. By means of this approach, the lag number of antecedent discharge is determined by auto-correlation function (ACF) and the lag number of antecedent rainfall is determined by cross-correlation function (CCF).

Multi Layer Perception (MLP)-Artificial Neural Networks (ANN)

The MLP-ANNs used in this study is a feed forward neural network which generally consists of three layers viz; input, hidden and output layers. Output y at time T is calculated from the function (8):

$$y\left(\frac{T}{\theta}\right) = g(\delta(T), \theta) \quad (8)$$

Where δ (T) is input vector (regressor) and θ represent activation function, sigmoid function for MLP.

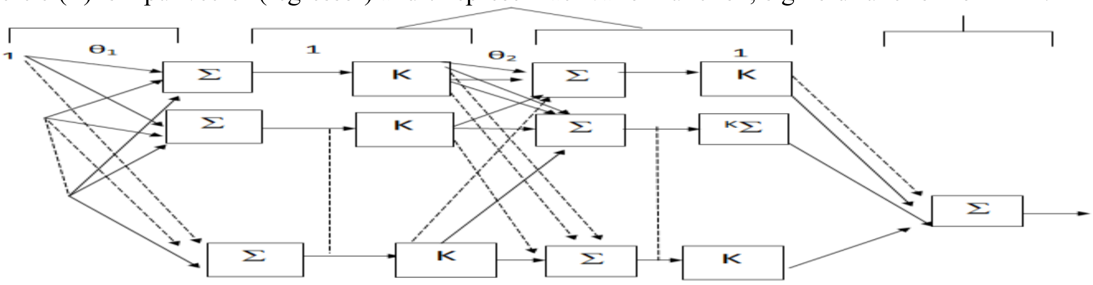


Figure 4: Feed Forward Multi Layer Perception

Improvement Error Back-propagation Algorithm

The error back-propagation algorithm is widely used to adjust the weight matrix and the biases matrix of the networks through an interactive training procedure for the purpose of revealing the relationship between the predictors and the predictands. From an initial random weight matrix and bias matrix, this algorithm searches the "weight space" by using gradient decent method to minimize the overall error between the simulated outputs and the observed values. A method with self-adaptive learning rate and self-adaptive momentum coefficient is used here to accelerate the training process and to prevent the algorithm from converging at a local minimum. The learning rate α and the momentum coefficient β are automatically adjusted as follows (Sudhear, 2002)

$$e_r^{(K)} = \frac{\Delta E(K)}{EK} = \frac{E(K) - E(K-1)}{E(K)} \quad (9)$$

Levenbergy-Marquardt Back-propagation Training

In prediction context, MLP-ANN training consists of providing input-output examples to the network, and minimizing the objective function (i.e. error function) using either a first order or a second order optimization method. This so-called supervised training can be formulated as one of minimizing as function of the weight, the sum of the nonlinear least squares between the observed and the predicted outputs, defined by:

$$E = \frac{1}{2}\sum_{p=1}^{n}\sum_{k=1}^{m}\left(y_{pk} - \hat{y}_{pk}\right)^2 \quad (10)$$

Where *n* is the number of patterns (observations) and *m* the total output units, y represents the observed response ("target output") and $\hat{y}$ the model response ("predicted output"). In the case of one output unit ($m = 1$) reduces to

$$E = \frac{1}{2}\sum_{p=1}^{n}\left(y_p - \hat{y}_p\right)^2 \quad (11)$$

Which is the usual function that is minimized in least squares regression. In the BP training, minimization of *E* is attempted using the steepest descent method and computing the gradient of the error function by applying the chain rule on the hidden layers of the MLP-ANN (Rumelhart et al., 1986). Consider a typical multi layer MLP-ANN (Figure 4) whose hidden layer contains *M* neurons. The network is based on the following equations:

$$net_{pj} = \sum_{i=1}^{N} W_{ji}\, x_{pi} + W_{jo} \quad (12)$$

$$g\left(net_{pj}\right) = \frac{1}{1 + e^{-net_{pj}}} \quad (13)$$

where net_{pj} is the weighted inputs into the jth hidden unit, N the total number of input nodes, W_{ji} the weight from input unit i to the hidden unit j, x_{pi} a value of the ith input for pattern p, W_{jo} the threshold (or bias) for neuron j, and $g(net_{pj})$ the jth neuron's activation function assuming that g() is the sigmoid function. Note that the input units do not perform operation on the information but simply pass it onto the hidden node. The output unit receives a net input of

$$net_{pj} = \sum_{j=1}^{M} W_{kj}\, g(net_{pj}) + W_{ko} \quad (14)$$

$$\hat{y}_{pk} = g(net_{pk}) \quad (15)$$

where M is the number of hidden units, W_{kj} represents the weight connecting the hidden node j to the output k, W_{ko} is the threshold value for neuron k, and $\hat{y}_{pk}$the kth's predicted output. Recall that the ultimate goal of the network training is to find the set of weights W_{ji} connecting the input units i to the hidden units j and W_{kj} connecting the hidden units j to output k, that minimize the objective function (Eq. (10)). Since Eq. (3.10) is not an explicit function of the weight in the hidden layer, the first partial derivates of E are evaluated with respect to the weights using the chain rule, and the weights are moved in the steepest-descent direction. This can be represented mathematically as

$$\Delta W_{kj} = -\eta \frac{\partial E}{\partial W_{kj}} \quad (16)$$

where η is the learning rate which simply scales the step size. The usual approach in BP training consists in choosing η according to the relation $0 < \eta < 1$. From Eq. (16), it is straightforward that BP can suffer from the inherent slowness and the local search nature of first order optimization method. However, BP remains the most widely used supervised training method for MLP-ANN because of the available remedies to its drawbacks. In all, second order nonlinear optimization techniques are usually faster and more reliable than any BP variant (Masters, 1995). Therefore, LMBP for MLP-ANN was used for data training. The LMBP uses the approximate Hessian matrix (second derivatives of E) in the weight update procedure as follows:

$$\Delta W_{kj} = -[H + \mu I]^{-1} J^T r \quad (17)$$

where r is the residual error vector, μ a variable small scalar which controls the learning process, $J = \nabla E$ is the Jacobian matrix, and $H = J^T$ denotes the approximate Hessian matrix usually written as $\nabla^2 E = 2J^T J$. In practice, LMBP is faster and finds better optima for a variety of problems than do the other usually methods (Hagan and Menhaj, 1994).

Design of MLP-ANN Architecture

The number of predictors and predicands specified the number of neurons in the input and output layers respectively. An experiment with trial-and-error measure, recommended as the best strategy by Shamseldin (1997) is used to determine the number of neurons in the hidden layer. In general, the architecture of multi-layer MLP-ANN can have many layers where a layer represents a set of parallel processing units (nodes). The three-layer FNN. (Figure 4) used in this study contains only one intermediate (hidden) layer. MLP-ANN can have more than one hidden layer; however theoretical works have shown that a single hidden layer is sufficient for ANNs to approximate any complex nonlinear function (Cybenko, 1989; Horinik et al., 1989). Indeed many experimental results seem to confirm that one hidden layer may be enough for most forecasting problems (Zhang et al., 1988; Coulibaly et al., 1999). Therefore, in the study, one hidden layer FNN is used. It is the hidden layer nodes that allow the network to detect and capture the relevant pattern(s) in the data, and to perform complex nonlinear mapping between the input and the output variables. The sole role of the input layer of nodes is to relay the external inputs to the neurons of the hidden layer. Hence, the number of input nodes corresponds to the number of input variables (Figure 4). The outputs of the hidden layer are passed to the last (or output) layer which provides the final output of the network. The network ability to learn from examples and to generalize depends on the number of hidden nodes. A too small network (i.e. with every few hidden nodes) will have difficulty learning the data, while a too complex network tends to overfit the training samples and thus has a poor generalization capability. Therefore, in this research, the trial-and-error method commonly used for network design was used. The training algorithm used is hereafter presented.

Performance Assessment of Rainfall-Inflow Model

A number of error measures (Dawson et al; 2007; Legates and Mccabe, 1999) have been developed to assess the goodness of fit performance of hydrological forecasting models but no standard has been specified since each measure can just assess one or two aspects of the runoff characteristic. Five commonly used error measures, therefore, are to be employed in this study to make the evaluation of the forecasts. They are coefficient of efficiency (CE), the mean absolute error (MAE), the squared relative error (MSRE), the relative volume error (RVE) and the coefficient of determination (R^2), respectively defined as follows:

$$CE = 1 - \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)^2}{\sum_{i=1}^{n}(Q_i - \bar{Q})^2} \quad (18)$$

$$MAE = \frac{\sum_{i=1}^{n}/Q_i - \hat{Q}_i/}{n} \quad (19)$$

$$MSRE = \frac{\sum_{i=1}^{n}\frac{(Q_i - \hat{Q}_i)^2}{Q_i^2}}{n} \quad (20)$$

$$RVE = \frac{\sum_{i=1}^{n}(Q_i - \hat{Q}_i)}{\sum_{i=1}^{n}(Q_i)} \quad (21)$$

$$R^2 = \left[\frac{\sum_{i=1}^{n}(Q_i - \bar{Q})(\hat{Q}_i - \hat{\bar{Q}})}{\sqrt{\sum_{i=1}^{n}(Q_i - \bar{Q})^2(\hat{Q}_i - \hat{\bar{Q}})^2}}\right]^2 \quad (22)$$

Where, Q_i is the observed discharge, $\hat{Q}_i$ is the simulated discharge, $\bar{Q}$ is the mean of the observed discharges, $\hat{\bar{Q}}$ is the mean of the simulated discharges and n is the length of the observed/simulated series.

The CE (Nash and Sutcliffe., 1970), which ranges from - ∞ to 1, describes the relationship between the modelled value and the mean of the observed data. A coefficient of one (CE = 1) means that the model performs a perfect matching to the observed data, a coefficient (CE = 0) indicates that the model result just has the equal mean value with the observed data, while coefficient of negative value (CE<0) shows that the model performs worse than using the observed mean. The MAE, which ranges from 0 to + ∞, is used to measure how close forecasts are to the eventual outcomes. Theoretically, a coefficient of zero (MAE = 0) means the best model with a perfect performance. The MSRE, which ranges from 0 to + ∞, can provide a balanced evaluation of the goodness of fit of the model as it is more sensitive to the larger relative errors caused by the low value and the best coefficient will be zero (MSRE = 0). The RVE, gives the relative bias of the overall water balance of the model and the best coefficient will be zero (RVE = 0) The $R^{2,}$ which ranges from 0 to 1, is a statistical measure of how well the regression line close to the observed data and coefficient of one (R^2=1) indicates that the regression line perfectly fits the observed data.

III. RESULTS AND DISCUSSION

Tables 1, 2 and 3 show the R^2 (coefficient of determination) tests for the analysis to determine the number of input nodes, hidden nodes and the training epochs for MLP network trained using back propagation algorithm (MLP-BP); From equations 9, the learning rate, α and the momentum coefficient, β were automatically adjusted according to Sudhear (2002) to yield: α= 0.001 and β = 0.85. The results in table 1 were produced by assigning hidden node for the neural networks model and the number of input nodes was varied to identify the best input node required by the neural network. It is clear that the best R^2 tests were produced with one input node, i.e. when only the last inflow lag is used. By assigning input node to one and changing the number of hidden nodes, the results in table 2 were obtained. The results indicated that the best number of hidden nodes for MLP-BP is 7.

Table 1: Variation of R^2 Tests with Different Input Nodes

No. Of Input Nodes	MLP-BP
1	0.9290
2	0.3927
3	-0.3635
4	-0.4949
5	-0.4949

The analysis to find the adequate training epochs was carried out and the results are shown in table 4.3. The results suggested that the adequate training epoch is 1000 for MLP-BP.

Table 2: Variation of R^2 Tests with Different Hidden Nodes

No. of hidden nodes	MLP-BP
1	-1.5959
2	0.8946
3	0.9290
4	0.9536
5	0.9521
6	0.9508
7	0.9570
8	0.9543
9	0.9507
10	0.9296
11	-
12	-
13	-
14	-
15	-
20	-

In order to test the generalization properties of the neural networks models, R^2 tests for multi-step-ahead (MSA) forecasting of the inflow were calculated. The neural network models were trained using the following structure:

MLP-BP: input node = 1, Hidden Nodes = 7, Training Epochs = 1000

Table 3: Variation of R^2 Tests with the Number of Training Epochs

No. of epoch	MLP-BP
1	-0.6300
2	-0.6278
3	-0.6233
4	-0.6129
5	-0.5847
6	-0.5123
7	-0.3538
8	-0.1210
9	0.0854
10	0.2498
12	0.5184
14	0.7317
16	0.8567
18	0.8996
20	0.9242
100	0.9570
1000	0.9706
1500	-
2000	-
3000	-

R^2 tests for lead-time from 1-day up to 6-day of the inflow were calculated over both the training and independent data sets as shown in table 4. The results for training data set indicate that MLP-BP gave good R^2 tests up to 4-day, 5-day and 6-day lead-time respectively, where their R^2 test values are about 0.8. The results for independent data set in table 4 showed that MLP-BP gave good R^2 tests up to 4-day, 5-day and 6-day ahead, respectively.

Another aspect that needs to be considered for on-line modelling is the complexity of the model. For on-line modelling and forecasting, all the calculation for parameter estimation or adjustment must be carried out within the sampling time. Hence, the number of adjustable parameters should be as small as possible. For MLP network, the adjustable parameter compose of connection weights between input nodes and hidden nodes, connection weights between hidden nodes and output node, and also the threshold in hidden nodes. Therefore, for one output network, the number of adjustable parameters for the network can be calculated using the following formula (Finnoff et al., 1993):

$$NAP_{MLP} = n_i \; x \; n_h + 2n_h = (n_i + 2)n_h \qquad (23)$$

Where *NAP*, n_i and n_h are the short form for number of adjustable parameters, number of input nodes and number of hidden nodes respectively. Based on the formula and the previously determined structure for the networks, the number of adjustable parameters for MLP networks was 16.

Table 4: R^2 Tests Calculated over both the Training and Independent Data Sets

Lead-time (days)	Training data set	Independent data set
	MLP-BP	MLP-BP
1	0.9888	0.9706
2	0.9771	0.9342
3	0.9491	0.8899
4	0.9098	0.7955
5	0.8672	0.7232
6	0.8138	0.6311

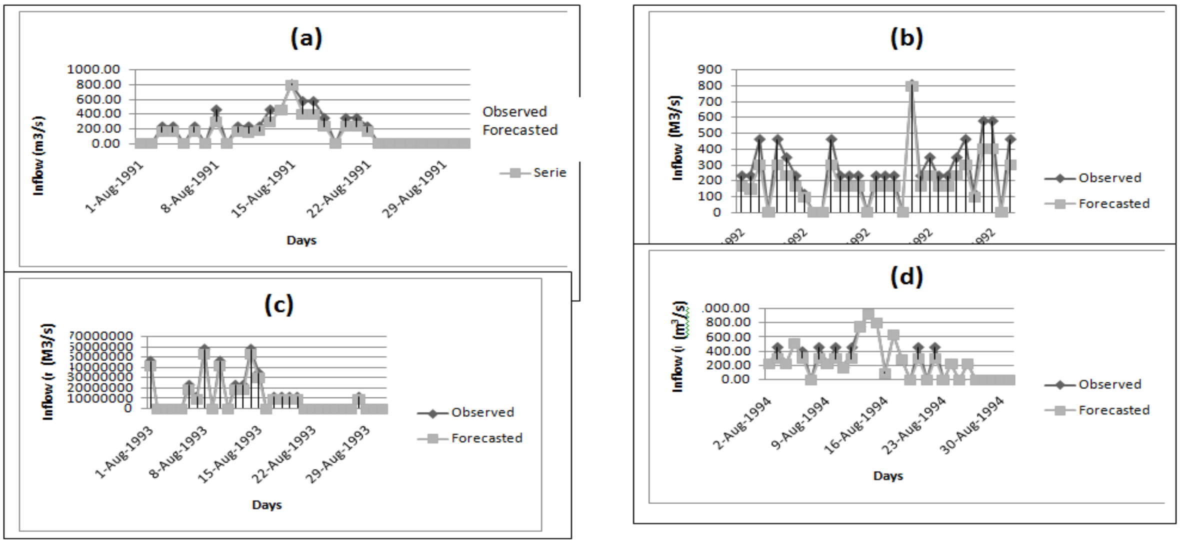
Figure 5 shows the relevant scatter diagrams of simulated discharge versus observed discharge of the Model in calibration (1991- 1998) for the months of August. Figure 6 presents the simulated results from the model synchronously given with the observed records in each figure for the months of August (1991-1998) during calibration. The simulated curves in Figures 6 and 7 clearly indicate that not only the rising trends and the falling trends in the hydrograph are picked up by the model but also excellent goodness of fits is achieved.

Results of Model Performance Assessment

The perfomance of the model was evalauted for the training period (1991-1998). From equations 16 to 20,

$$CE = 9.91\,X\,10^{-1}\,;\ MAE = 9.679\,X\,10^{-6};\ MSRE = 2.49\,X\,10^{-7};\ RVE = 3.840\,X\,10^{-6}\ and\ R^2 = (0.991)^2$$

Satisfactory forecasting is obtained in this study since the CE and R^2 are sufficiently high and close to 1, and the MSRE and RVE are adequately low and approximates to 0. The measures MAE of calibration and validation are far less than the relevant mean value of the observed data. The high scores of CE and R^2 indicate that all the models present the "best" performance according to the standard given by Dawson et al. (2007). The statistic result of error measures of the validation are as much as that of the calibration and both of them are encouraging. This outcome implies that the training procedures are successful without "overtraining" or "local minimum" and the proposed models have powerful generalization abilities for out-of-sample forecasting. Numerous MLP-BP-ANN structures tested in order to determine the optimum number of hidden layers and the number of nodes in each.



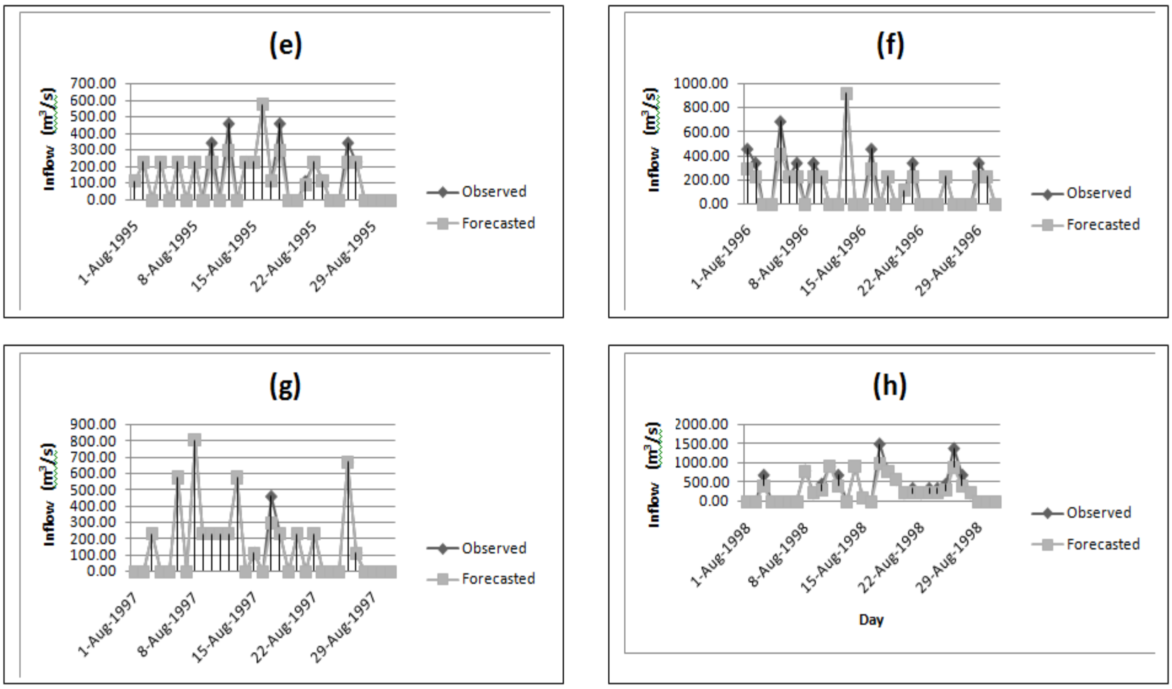


Figure 5: Forecasted and Observed Inflow Hydrographs of (a) August 1991(b) August 1992 (c) August 1993 (d) August 1994 (e) August 1995 (f) August 1996 (g) August 1997 (h) August 1998 in Model Calibration

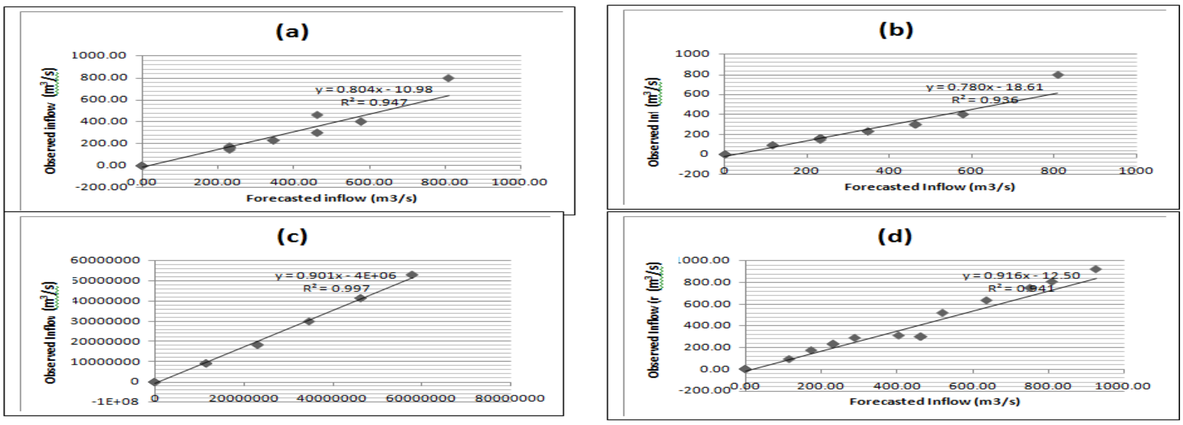


Figure 7: Forecasted and Observed Inflow Hydrographs of (a) August 1999 (b) August 2000 (c) August 2001 in Validation

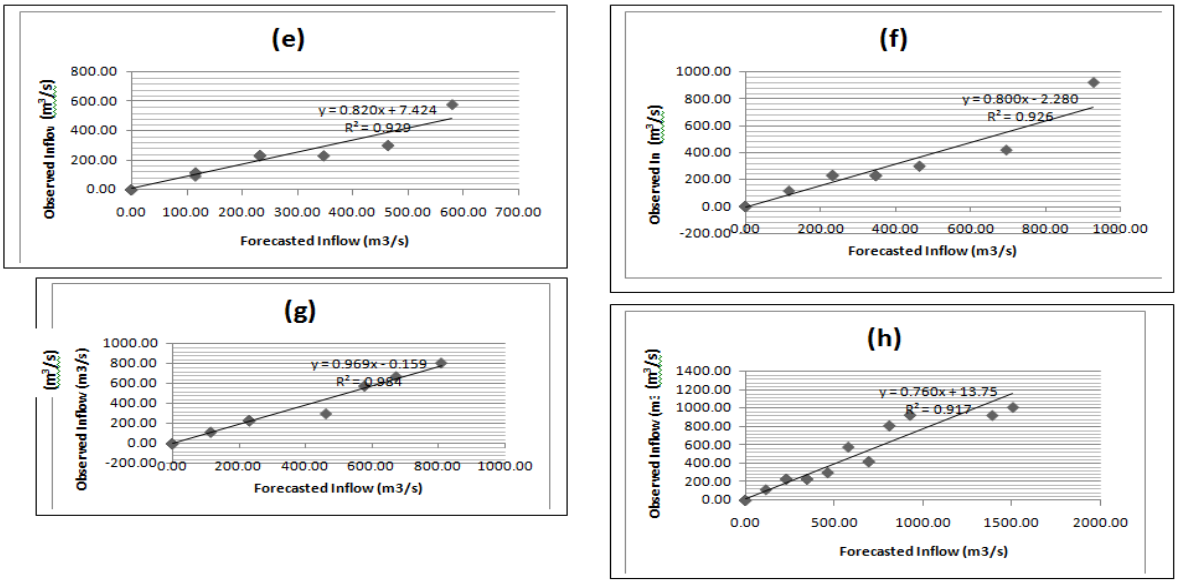


Figure 8: Scatter Plots of Observed and Simulated Discharges of(a)August 1999(b) August 2000(c) August 2001 in Validation

IV. CONCLUSION

In this work, Multilayer Perceptron Back Propagation Artificial Neural Network (MLP-BP-ANN) models were developed for forecasting daily inflow values into Dadin-Kowa reservoir. The Artificial Neural Network approach becomes more explicit and can be adopted for any reservoir daily inflow forecasting. The experimental results indicate that these models can extend the forecasting lead-times with a satisfactory goodness of fit.

As regards the accuracy, all the models provided good accuracy for short time horizon forecast which however decreased when longer time horizons were considered and this was particularly true for the rising phase of the flood wave where a systematic underestimation was observed. This temporal limit is coherent with that detected by other authors using similar data-driven models applied to basins with similar extension to that considered in this study (Solomatine and Dulal, 2003), and this limit is certainly due to the fact that no information or forecast of rainfall is considered available within the time spell ahead with respect to the instant when the forecast is performed.

REFERENCES

[1.] **Abrahart, R.J. and See L. (1998).** Multi-model Data Fusion for River-flow Forecasting: An Evaluation of Six Alternative Methods Based on Two Contrasting Constraints. Hydrol. Earth Syst Sci. 6, 655-670.

[2.] **Burke, L.I. (1991).** Clustering Characterization of Adaptive Resonance. Neural Network. 4(4), 485-491.

[3.] **Coulibaly, P., Anctil, F. and Bobee, B. (1999).** Prevision Hydrologique par

[4.] **Cybenko, G. (1989).** Approximation by Superposition of a Sigmoid Function. Math. Control Signals. 2, 303-314.

[5.] **Dadin-Kowa Local Government Area. (1999).** Potrait of Dadin-Kowa L.G.A, Gombe State Nigeria.

[6.] **Dandy, G. and Mainer, H. (1996).** Use of Artificial Neural Network for Real Time Forecasting of Water Quality. Proc. Int. Conf. Water Resource. Environ Res. Vol II October 29-31. Kyoto, Japan.

[7.] **Dawson, C.W., Abrahart R.J, and See, L.M. (2007).** Hydrotest : A web-based Toolbox of Evaluation Metrics for the Standardized Assessment of Hydrological Forecasts, Environ Model, Soft Ware 22, 1034-1052.

[8.] **Finnoff, W., Hergert, F. and Zimmermann, H.C. (1993).** Improving Model Selection by Non-convergent Methods. Neural Networks 6, 771-783.

[9.] **French, M.N., Krajaswski, W.F. and Cuykendall, R.R. (1992).** Rainfall Forecasting in Space and Time, using a Neural Network. Journal of Hydrology 137, 1-31.

[10.] **Hagan, K.T. and Menhaj, M.B. (1994).**Training Feed forward Networks with the Marguardt Algorithm. IEEE Trans-Neural Networks 5(6), 989-993

[11.] **Hall, M. J. and Minns, A. W. (1993).** Rainfall-runoff modelling. As a Problem in Artificial Intelligence: Experience with Neural Network. Proc. 4th National Hydrology Simposium. British Hydrological Society, Gidiff. pp 5.51-5.57.

[12.] **Hornik, K., Stinchcombe, M. and White, H. (1989).** Multilayer Feedforward Networks are Universal Approximators. Neural Networks 2(5), 359-366

[13.] **Hsu, K. L., Gupta H. U. and Soroshian, S. (1995).** Artifical Neural Network Modelling of Rainfall- runoff Process. Water Resour 31 (10), 2517-2530.

[14.] **Ibeje, A.O., Agunwamba, J. C. and Okoro, B.C. (2012).** Allocation of Water Resources in Semi-Arid Regions of Northern Nigeria (Poster Presentation). In : Integrated Water Resource Management: Hartwig Steusloff, ed. Conference Proceedings at

International conference Karlsruche, IWRM Karlsruche 2012, 12-22 November, 2012 at Karlsruche Convention Centre, Germany. ISBN 978-3-8396-0478-6.

[15.] **Legatees, D.R. and McCabe, G.J. (1999).**Evaluating the Use of "Goodness of Fit" Measures in Hydrologic and Hydroclimatic Model Validation. Water Resour.Res., 35, 233-241.

[16.] **LSWS. (1999).** Internet page of LUND SPACE WEATHER CENTER (LSWS). [On-Line] available. http://nastol.astro . in.se/._henrik/neuralnet3-html.

[17.] **Minns, M. (1998).** Metro Deterministc Use of Artificial Neural Networks for Prediction of Water Quality Parameters. Water Resour Res 32 (4), 1013-1022

[18.] **Morgan, S. (2000).** Morgan Stanley Wilter Research. http://www.Msdw.com/search/index html. Accessed: Jan 2011

[19.] **Nash. J.E and Sutcliffe, J.V. (1970).** River Forecasting through Conceptual Models Part 1-A Discussion of principles. Hydrol., 10, 282-290

[20.] **Nussbaum, P., Schreiba, T. and Schaffrath, C. (1996).** Nonlinear Time Sequence Analysis. Int. J. Bifurcation Chao.I(3), 521-547.

[21.] **Rumelhert, D. E, McClelland, J. L. and The PDP Research Group. (1986).** Parallel Recognition in Modern Computers. In Proceeding Explorations in the Micro Structure of Cognition. Vol. foundation, MIT press/Bradford books Cambridge mass.

[22.] **Salazar J. E., Caravajah, L. F., Mesa, O. J., and Smith, R. (1998).** Modelos de Prediccion de Caudales Mensuales Considerando Anomalia Climaticas. Ponencia XVIII Congresso Latinoam-De Hidraulica Oaxaca Mexico. pp 525-533.

[23.] **Shamseldin, A.Y. (1997).** Application of A Neural Networks Technique to Rainfall-Runoff Modelling. J. Hydrol., 199, pp 272-294

[24.] **Solomatine, D. P., and Dulal, K. N. (2003).** Model Trees as An Alternative to Neural Networks in Rainfall-Runoff Modelling. Hydrol. Sci J., 48(3), 399-411. Technical Paper No. 463, World Bank, Washington DC.

[25.] **Sudheer, K. B., Gosain, A.K. and Ramasastri, K.S. (2002).** A Data-driven Algorithm for Constructing Artificial Neural Network Rainfall-runoff Models. Hydrol. Process., 16, 1325-1330

[26.] **Zealand, C.M, Burn. D. H. and Simononic, S.P (1999).** Short-term Streamflow Forecasting using Artificial Neural Networks. Journal of Hydrology 214, 32-48.

[27.] **Zhang G., Patuwo, B.E. and Hu, M.Y. (1998).** Forecasting with Artificial neural Networks: the State of the Art. Int. J. Forecasting 14, 35-62.

[28.] **Zhu, M.L. and Fugita, M. (1994).** Comparison Between Fuzzy Reasoning and Neural Network Methods for Forecasting Runoff Discharge. Journal of Hydrolscience Hydraulic Engerng 12(12), 131-141.

[29.] **Zurada, R. J. (1992).** The Xinanjiang Model Applied in China. Journal of Hydrology. 13, 371-381.