Nigerian Journal of Water Resources. Vol. 4, No.1, December 2018.

SELECTION OF PARAMETERS USING ARTIFICIAL NEURAL NETWORK FOR WATER LEVEL PREDICTION IN A RESERVOIR

BELLO SAHEED A. AND IBRAHIM HUSSEIN
National Water Resources Institute, Kaduna.
Correspondence Author: KunleBello24@gmail.com, 08188691365

ABSTRACT
Accurate prediction of water levels in dams is very important in planning adequate water supply and power generation from dam to provide sufficient water storage during the critical periods. Effective power planning helps in ensuring steady supply of electric power to consumers to boost industrial activities and water supply to the community. The aim of this study is to develop artificial neural network models for predicting water levels at Dadin Kowa Dam, which is located 5 kilometers north of the village of DadinKowa on River Gongola in Gombe state. It involves taking of a ten-year record of the daily water levels at the dam from 2007 to 2016. The daily water level data were used to develop five neural network models. The results show that the prediction accuracy of the neural network models increased with increasing input. The Five-input layers neural network model had the lowest relative error (MODEL ERR0R: 0.00165582, RMSE: 0.06671, MARE: 0.000176) while the three-input layers model had the highest relative error (MODEL ERR0R: 2.72107, RMSE: 26087.445, MARE: 3.1272099). The neural network models which involve little mathematics were much simpler to build. The developed models will be very useful in water-use planning for irrigation, municipal uses and predicting power loads and management of power generation. Timely prediction can also help in disaster monitoring, response and control of floods in Nigeria

KEYWORDS: *Artificial neural network, Water level, Modeling, Feed forward error, Back propagation, RMSE, MARE*

INTRODUCTION

In every reservoir, the volume of water contained in it is well known, and by level monitoring, the flow through the reservoir and the extraction rate of water can be controlled to maintain a stable water supply. The demand for water will cause the water level in reservoirs to fall and level monitoring can control pumps to refill as required. Monitoring not only helps to prevent the service reservoir from overflowing but also from running empty and raising alarms if there is a failure in the pump control. At present, two main approaches are employed in hydrological prediction. The first approach is based on mathematical modeling. It models the physical dynamics between the principal interacting components of the hydrological system. In general, a rainfall-runoff model is used to transform the point values of rainfall, evaporation, and flow data into hydrograph predictions by considering the spatial variation in storage capacity. A hydraulic channel flow routing model is then used to calculate flow. An example of this type of deterministic modeling is River Flow Forecasting (RFFS), which is a large-scale operational system currently employed by the Outer River Catchment [Moore et al., 1994]. The second approach is based on modeling the statistical relationship between the hydrologic input and output, without explicitly considering the relationships that exist among the involved physical processes. Examples of stochastic models used in hydrology are autoregressive moving average models (ARMA) [Box and Jenkins, 1976] and the Markov method [Yakowitz, 1985].

For prediction of water level in a dam under scenarios of interest, different deterministic models have been attempted in the past. Advancement in Artificial Intelligence and computational Intelligence have led to the various techniques adopted in the prediction and modeling of the water quality. These techniques involves the use of Artificial Neural Network(ANN) ([Gustilo and Dadios, 2011; Miao, 2010]), Particle Swarm Optimization [Deng et al., 2006], the use of Genetic Algorithm, Fuzzy Logic Control [Chang and Xinrong,2013], Gray wolf Optimization[Sweidan et al.,2013].ANNs provide a quick and flexible means of creating models for estimation of stream water quality. In recent years ANNs have shown exceptional performance as regression tools, especially when used for pattern recognition and function estimation. In addition, there are many advantageous characteristics of ANN approach to problem solving viz.: (1) application of a neural network does not require a priori knowledge the underlying process; (2) one may not recognize all the existing complex relationships between various aspects of the process under investigation; (3) a standard optimization approach or statistical model provides a solution only when allowed to run to completion whereas a neural network always converges to an optimal (sub-optimal) solution condition and; (4) neither constraints nor an a priori solution structure is necessarily assumed or strictly enforced in the ANN development [Nan et al., 2006; Schtz et al., 2015]. These characteristics render ANNs to be very suitable tools for handling various hydrological modeling problems. However, the number of uses for ANNs is increasing rapidly and in recent years they have been successfully used for the prediction of economic, water resources, water quality and hydrologic time series [O¨zgu¨r. Kisi, 2011; NBCBN, 2005].

Methods and Materials

An ANN is a computing system made up of a highly interconnected set of simple information processing elements, analogous to a neuron, called units. The neuron collects inputs from both a single and multiple sources and produces output in accordance with a predetermined non-linear function. An ANN model is created by interconnection of many of the neurons in a known configuration. The primary elements characterizing the neural network are the distributed representation of information, local operations and non-linear processing. The learning process or training forms the interconnection between neurons and is accomplished by using known inputs and outputs, and presenting these to the ANN in some ordered manner. The strength of these interconnections is adjusted using an error convergence technique so that a desired output will be produced for a known input pattern [Lachtermacher and Fuller, 1994; Rumelhart et al, 1986].

These weights are updated or modified interactively using the generalized delta rule [ASCE Task Committee, 2000].

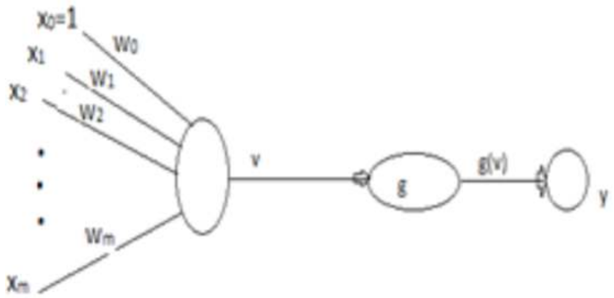

Fig 1. The proposed model

The layers consist of neurons whose output is a function of weights and activation functions in the neurons, and is defined by:

$$y = \varphi\left(\sum_{i=0}^{m} \left(w_{ji} x_i + b_j\right)\right)$$

Where φ *is the activation function of the layer*
w *is the interconnecting weight*
x *is the input to the neuron*
b *is the connecting bias*
j *is the number of neurons in the layer*
i *is the number of input*

In order to predict the accuracy of the water level, two error measures are used to compare the ANNs output with observed values: Root Mean Square Error (RMSE), and Mean Absolute Relative Error (MARE). According to [Karunanithi et al, 1994], Root Mean Square Error (RMSE) provide a good measure of the goodness of fit at high flows, whilst Mean Absolute Relative Error (MARE) provide a more balanced perspective of the goodness of fit at moderate flows. They are calculated as follows:

$$RMSE = \sqrt{\frac{1}{p}\sum_{i=1}^{p}(\theta_i - \theta_m)^2}$$

$$MARE = \frac{1}{p}\sum_{i=1}^{p}\left(\left|\frac{\theta_i - \theta_m}{\theta_i}\right|\right)$$

θ_i = predicted value, and θ_m = measured value.
Where p is the total number of input

Study Area

The multipurpose earth fill Dadinkowa dam is located some 5km. North of Village of Dadin Kowa on river Gongola in Gombe state. Dadin Kowa dam studies and investigation were carried out 2003 to 2016. The maximum flood level is 249m, maximum supply level is 247meters and the minimum supply level is 239meters, the surface area of the reservoir is 300km^2 with a live storage of 1.77billion cubic meters. The 1:10,000 year peak in-flow flood is 3,160M^3/Sec. and the peak outflow is 1,110M^3/Sec. the total catchment area of the Gongola River is approximately 56,000 square kilometers, 58.5% of which lies upstream of the dam.

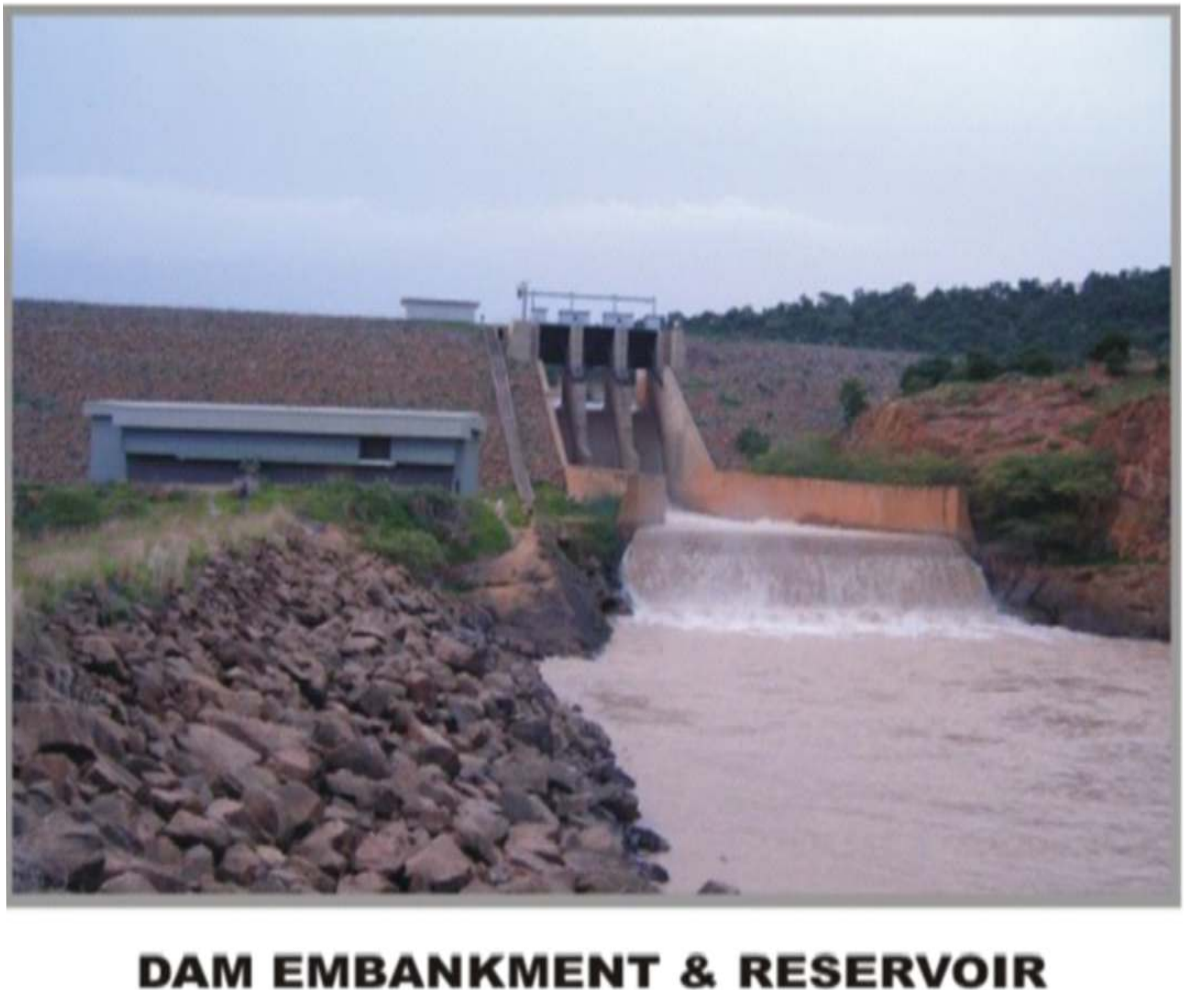

Figure 1 Dadin Kowa Dam. Accessed from compendium of Nigerian Dam, 2007

The importance of predicting water level of Dadin Kowa dam is to plan as follows:

- exploitation of water resources in future;
- water shortage determination;
- save water in excess of usage;
- determining consumption patterns and generating energy; and
- flood control.

These predictions can be made for different durations; short time prediction for duration less than one week, middle time for monthly and seasonal prediction and long time for annual prediction

.Result and materials

In the present study, the input data for a variable x were standardized by the ANN model. To simulate the water level, ANNs was developed using the Matrix Laboratory (MATLAB) software based on back propagation algorithm. The models stopped at six days because the study was based on short time prediction for duration less than or equal to one week.

Table 1: Arrangement of daily water level into five days

x1	**x2**	**x2**	**x4**	**X5**
239.906	239.886	239.876	239.866	239.856
239.846	239.836	239.826	239.816	239.806
239.796	239.776	239.776	239.766	239.756
239.746	239.746	239.736	239.726	239.716
239.706	239.686	239.676	239.666	239.656
239.646	239.636	239.626	239.616	239.606
239.596	239.596	239.586	239.586	239.576
239.576	239.666	239.666	239.556	239.546
239.546	239.536	239.526	239.526	239.466
239.456	239.456	239.446	239.446	239.436
239.426	239.416	239.406	239.396	239.396
239.376	239.366	239.356	239.356	239.346
239.336	239.226	239.216	239.206	239.196
239.196	239.186	239.186	239.176	239.166
239.156	239.146	239.136	239.126	239.116
239.106	239.096	239.066	239.056	239.046
239.036	239.036	239.026	239.016	239.006
239.986	238.976	238.966	238.956	238.946
238.926	238.906	238.896	238.876	238.866
238.846	238.836	238.816	238.806	238.796
238.786	238.776	238.766	238.756	238.746
238.726	238.726	238.716	238.716	238.706
238.706	238.696	238.696	238.686	238.676
238.676	238.666	238.666	238.656	238.646
238.646	238.636	238.636	238.626	238.626
238.616	238.616	238.606	238.606	238.596
238.646	238.626	238.616	238.606	238.606
238.596	238.586	238.576	238.576	238.566
238.566	238.556	238.556	238.546	238.546
238.536	238.526	238.516	238.516	238.506
238.496	238.496	238.486	238.486	238.516
238.516	238.526	238.556	238.556	238.557
238.606	238.606	238.596	238.586	238.576
238.596	238.576	238.596	238.596	238.656
238.706	238.806	238.846	238.876	238.886
238.906	238.996	239.026	239.006	239.106
239.166	239.206	239.226	239.246	239.266
239.286	239.306	239.366	239.386	239.406
239.506	239.546	239.726	239.756	239.816
239.906	239.926	239.946	240.066	240.106
240.146	240.206	240.306	240.406	240.586
240.666	240.726	240.846	240.966	241.026
241.056	241.086	241.146	241.226	241.406
241.586	242.506	242.666	242.906	243.106
243.306	243.466	243.506	243.626	243.706
243.826	243.966	244.006	244.206	244.306
244.466	244.586	244.706	244.826	245.006
245.166	245.306	245.586	245.866	246.476
246.916	247.306	247.556	247.606	247.616
247.601	247.566	247.506	247.411	247.336
247.296	247.156	247.056	246.976	246.866
246.751	246.626	246.496	246.372	246286
246.191	246.076	246.016	245.886	245.876
245.866	245.716	245.736	245.696	245.626
245.516	245.351	245.396	245.236	245.116

245.021	244.906	244.781	244.711	244.491
244.391	244.731	244.151	244.026	243.921
243.726	243.666	243.561	243.456	243.356
243.246	243.146	243.036	242.966	242.846
242.746	242.651	242.561	242.486	242.486
242.306	242.236	242.146	242.066	245.986
241.916	241.886	241.846	241.716	241.676
241.586	241.516	241.426	241.376	241.316
241.276	241.216	241.166	241.126	241.076
241.026	240.946	240.926	240.881	240.846
240.801	240.756	240.716	240.666	240.651
240.616	240.566	240.526	240.506	240.456
240.426	240.406	240.366	240.326	240.306
240.271	240.236	240.216	240.191	240.166
240.116	240.096	240.086	240.076	240.066
240.051	240.026	240.006	239.946	239.963
239.916	239.876	239.866	239.856	239.846

For ANN development, the daily records of water level were taken as number of combinations of input and target variables namely:

1. first day as input and second day as target;
2. first and second day as input and third day as target;
3. first, second, third day as input and forth day as target;
4. first, second, third, fourth day as input and fifth day as target;
5. first, second, third, fourth, fifth day as input and sixth day as target.

For the analysis, there are different ANN model for the cases (1), (2), (3), (4) and (5) respectively. The output layer had a single node corresponding to the inputs. The network was trained by;

```
net = newff(inputt,targett,10,{'tansig','purelin'},'trainlm','learngdm')
```

The comparative performance of various ANN models based on Model error, RMSE and MARE are given in Table 1. It can be seen from the Table that the values of Model error vary from 0.00165582 to 2.72107, RMSE from 0.06671 to 27924.064 and MARE from 0.000176 to 3.1272099.

As explained above, five types of ANN models have been developed with different combinations of data, i.e., (1) first day as input and second day as target (2) first and second day as input and third day as target (3) first, second, third day as input and forth day as target (4) first, second, third, fourth day as input and fifth day as target (5) first, second, third, fourth, fifth day as input and sixth day as target.

Performance of the five ANN models is different. The best performing ANN model is the MODEL5 with values of Model error, RMSE and MARE are 0.00165582, 0.06671 and 0.000176 respectively. MODEL5 has a total of 60 input variables consisting of 5 layers. The results indicate that the MODEL5 performances are best with number of input layers. The graphical results of the best performing ANN, i.e., MODEL5 are shown below in table and graphs which depict a good match between the observed and simulated results by ANN method.

Table 2: Comparative performance of various ANN models

MODEL	MODEL ERR0R	RMSE	MARE
MODEL1	0.152448	0.323569	0.000444
MODEL2	0.0279235	0.159613	0.000401
MODEL3	2.72107	26087.445	3.1272099
MODEL4	0.437375	27924.064	0.9944672
MODEL5	0.00165582	0.06671	0.000176

Table 2 shows the comparative performances of the ANN model. MODEL5 (Five-input layers) had the smallest model error, RMSE and MARE follow by MODEL2 (two-input layers) with a large variation between them. Therefore, by error estimation MODEL5 performed better than others.

MODEL2 also performed well but looking at the Figure 4 the predicted values are mostly below the actual values even at the peak compared to Figure 3 of MODEL5. The results show that MODEL2 does not put flood prevention into consideration which made it non-useful in this study. After observing the evaluation parameters, it is observed that Feed Forward Back propagation MODEL5 gives the best output having model error; 0.00165582, RMSE; 0.06671, MARE; 0.000176.

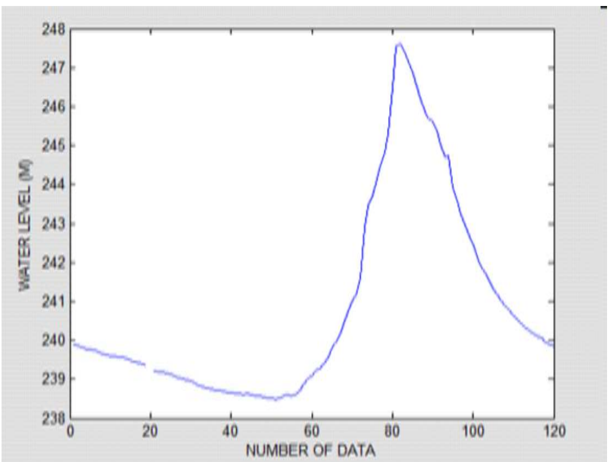

Figure 2: Sample Reservoir Levels for Training Data set

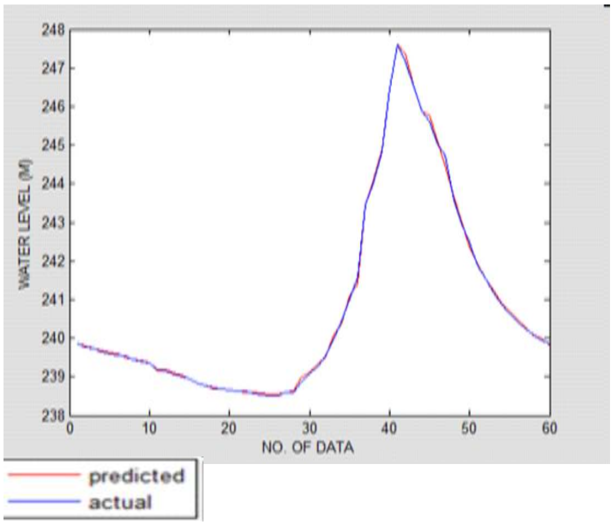

Figure 3: Actual and Predicted Water levels for MODEL 5 using ANN

The MODEL5 predicts the water levels that measure very close or near to the actual levels, as depicted in the form of graph in Figure 3. Less variation measure seen with the forecasted and that of the actual provides the closeness of predication to the actual values.

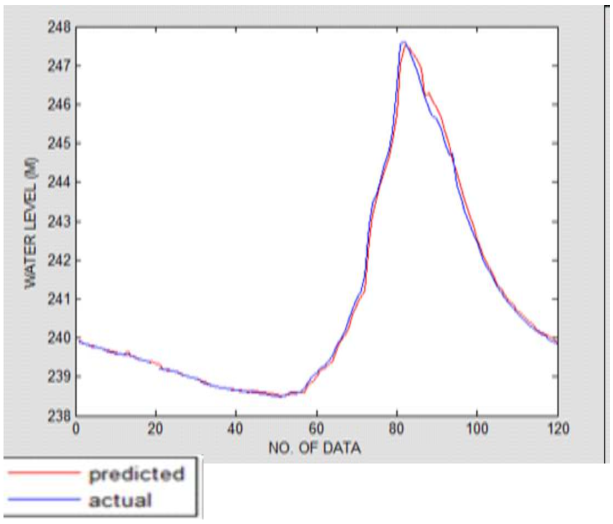

Figure 4: Actual and Predicted Water levels for MODEL 2 using ANN

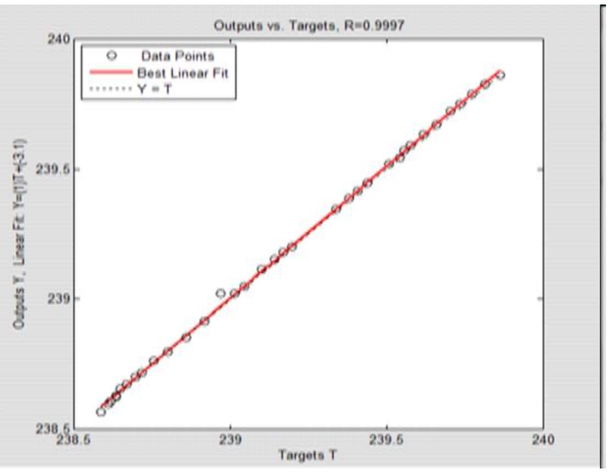

Figure 5: Correlation coefficient(R) between the outputs and Targets on MODEL5

The correlation coefficients that indicate the strength of the relationship between observed and predicted data are higher than 0.9 (maximum scale is 1)[Joorabchi et al, 2007]. The best result was predicted for MODEL5 with a correlation coefficient equals 0.9997 in Figure 5 and Correlation coefficient(R) of Training, Validation and test outputs against Targets in Figure 6 are 0.99949, 0.99996 and 0.99995 respectively. This shows the Correlation coefficient(R) for MODEL5 indicate strong strength of relationship between observed and predicted data.

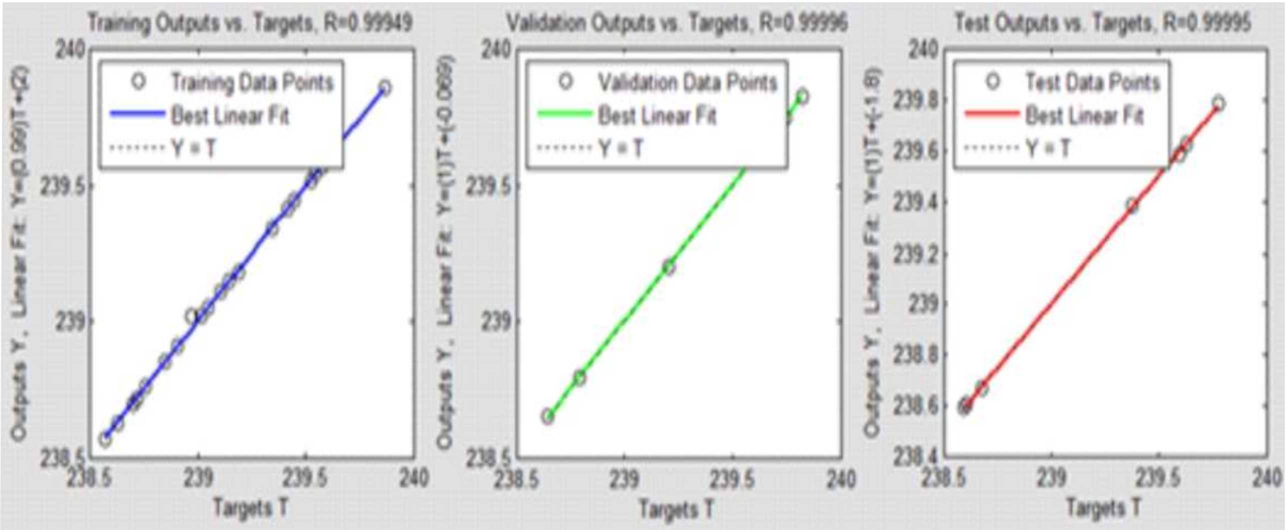

Figure 6: Correlation coefficient(R) of Training, Validation and test outputs and Targets on MODEL5

Conclusions

An artificial neural network model with a Back-propagation learning algorithm is adopted in this study to provide an effective and timely prediction of Water level in Dadin Kowa dam. This can help in water-use planning for irrigation, municipal uses and predicting power loads and management of power generation. Timely forecasting can also help in disaster monitoring, response and control of floods. Different network structures were compared and the performances were tested using MODEL ERROR, RMSE and MARE. It was found that the Five-layer input of ANN approach turned out to be an efficient approach and had better outcomes for predicting water level. A cross validation method was used to prevent the network from over-fitting. Results show the neural network with five-layer input provides a high accuracy prediction of water level in the reservoir. One of the advantages of the presented model compared to the ordinary numerical models is that it is not dependent on the initial and boundary conditions.

The ANNs have been successfully used in many hydrological studies and this was a motivating factor for its application to the present study. However, the input data should be consistent and the controlling factors should be the same for training and test data.

REFERENCES

ASCE Task Committee, (2000).Artificial neural network in hydrology. I: Preliminary concepts. Journal of Hydrologic Engineering, ASCE, 5(2), 115-123.

Box G.E.P, Jenkins G. M. (1976).Time Series Analysis: Forecasting control, Holden – Day, Oakland, California.

Chang, B., & Xinrong, Z. (2013). Aquaculture Monitoring System Based on Fuzzy-PID Algorithm and Intelligent Sensor Networks. Paper presented at the Cross Strait Quad-Regional Radio Science and Wireless Technology Conference (CSQRWC).

Deng, C., Wei, X., & Guo, L. (2006). Application of Neural Network Based PSO Algorithm in Prediction Model for Dissolved Oxygen in Fish Pond. Paper presented at the 6th World Congress on Intelligent Control and Automation, Dalian Chain.

Gustilo, R.C., &Dadios, E. (2011). Optimal control of prawn aquaculture water quality index using artificial neural networks. Paper presented at the 2011 IEEE 5th International Conference on Cybernetics and Intelligent Systems (CIS).

Joorabchi, A., Zhang, H. and Blumenstein, M., (2007). Application of artificial neural networks in flow discharge prediction for the Fitzroy River, Australia. Journal of Coastal Research, SI 50 (Proceedings of the 9th International Coastal Symposium), 287 – 291. Gold Coast, Australia, ISSN 0749.0208

Karunanithi, N. Grenney, W.J. Whitley, D. and Bovee, K. 1994: Neural networks for river flow prediction. Journal of Computing in Civil Engineering 8, 201–20.

Lachtermacher, G. and Fuller, J.D. (1994). Back propagation in hydrological time series forecasting. In Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Vol.3, Time Series Analysis in Hydrology and Environmental Engineering, pp. 229-242, edited by k.W.Hipel, A.I. Mcleod, U.S.Panu and V.P.Singh, Kluwer Acd., Norwell, Mass..

Miao, X., Deng, C., Li, X., Gao, Y., & He, D. (2010). A Hybrid Neural Network and Genetic Algorithm Model for Predicting Dissolved Oxygen in an Aquaculture Pond. Paper presented at the 2010 International Conference on Web Information Systems and Mining (WISM).

Moore, R.J., Jones, D.A., Black, K.B., Austin, R. M., Carrington, Tinnion, M., Akhondi, A., (1994). RFFS and HYRAD: Integrated System for Rainfall and River Flow Forecasting in Real-Time and their Application in Yorkshire. BHS Ossasional paper No. 4, 12.

Nan, L., Zetian, F., Ruimei, W., & Xiaoshuan, Z. (2006). Developing a Web-based Early Warning System for Fish Disease based on Water Quality Management. Paper presented at the 2006 1ST IEEE Conference on Industrial Electronics and Applications.

Nile Basin Capacity Building Network (NBCBN) (2005). Flood and Drought Forecasting and Early Warning Program.

O¨zgu¨r. Kisi (2011). A combined generalized regression neural network wavelet model for monthly stream flow prediction, KSCE J.Civil Eng. 15 (8) 1469–1479.

Rumelhart, De. E., Hinton, G. E. and Williams, R.J. (1986). Learning representation by back propagating errors. Nature, 323(9), 533-536.

Schtz, F., Lima, V. d., Eyng, E., & Bresolin, A. A. (2015). Simulation of the concentration of dissolved oxygen in river waters using artificial neural Networks. Paper presented at the 11th International Conference on Natural Computation (ICNC).

Sweidan, A. H., El-Bendary, N., Hassanien, A. E., &Elkarim, O. M. H. A. (2013). Water Quality Classification Approach based on Bio-inspired Gray Wolf Optimization. Paper presented at the 7th IEEE International Conference of Soft Computing and Pattern Recognition.

Yakowitz, S.J. (1985), Markov flow models and the flood warning problem, Water Resources 21 81–88.