Nonlinear Structure based Artificial Neural Computing for Upstream Flow Functional Models

Tanveer Ahmed Siddiqi,
Syed Inayatullah,
Syed Ahmad Hassan,
Saba Naz,
Department of Mathematics, University of Karachi, Karachi, Pakistan
Syed Tanweer Iqbal,
Institute of Space and Planetary Astrophysics,
University of Karachi, Karachi, Pakistan
Muhammad Ahsanuddin,
Department of Economics, University of Karachi, Karachi, Pakistan

Doi: 10.19044/esj.2019.v15n15p1 URL:http://dx.doi.org/10.19044/esj.2019.v15n15p1

Abstract

Most of the real world systems are nonlinear and complex and it is challenging to model these types of systems for analyzing and forecasting the hidden behaviour of the systems. In the paradigm of vague complex systems, data-based time series modeling approaches of intelligent systems showed its applicability for coping with the problems of hidden noise and dynamicity which are encapsulated in the data. Getting from nature is one of the humans' features and they are striving to produce the intellectual schemes by coping rare features of cognitions and intellect of the brain. In this paper, nonlinear autoregressive structure based modeling of the brain (i.e. Artificial Neural Network) is the aim of this study that suggest various Dynamic Neural Network (DNN) models by using time deferred autoregressive configurations, for the stream-flow of Sukkur barrage on lower Indus river basin. The suitability of the models for training, validation and testing stages, are evinced on assessment metrics which demonstrate the accuracy and sufficiency of the models which may be beneficial for water-resource management.

Keywords: Intelligent systems, Nonlinear autoregressive structure, Artificial neural network, Indus basin

Introduction

The most of the areas in Pakistan are facing the problem of deficiency or scarcity of water because of the aridness in the country due to a lesser amount of precipitation. Therefore, it is necessary to manage the existing water resources in the forms of dams or barrages so that it may fulfill the requirement of water consumption in the country. A time-ordered sequence of data values of a physical system made at equal intervals of time is called time series and represented by a set of discrete values $x_1, x_2, x_3, \dots \dots \dots$, etc. Generally, the time series based on measured values and typically fluctuated by noise, it normally contains a deterministic signal element and a random element showing the abrupt intervention that produces statistical variations about the deterministic values. These intelligent tools or techniques, including searching optimization, mathematical optimization, logic, and learning based statistical techniques (Luger, 2005), are also constructive to cope with the issues of dynamicity and noise hidden in the data or observations. In the last few spans, data-driven time series modeling or forecasting has developed a prevalent research area and has become dominated consideration for researchers or scientists. Data-driven time series modeling has recognized as a headed mechanism for characterizing and analysis of complex systems from real observed nonlinear or nonstationary data (Kantelhardt et al., 2002; Lai and Ye, 2003; Zhang et al., 2007). Many researchers have focused to employ data-driven time series modeling or analysis (Cherif, Cardot and Boné, 2011; Wan, Gong and Si, 2016; Wang, Wang and Liu, 2016; Hu, 2017; Ben Taieb et al., 2012) in which time series forecasting models exhibit a major role in usage point of view (Li et al., 2013; Bodyanskiy and Popov, 2006; Li and Song, 2008; Chen, Lai and Yeh, 2012; Lu et al., 2015; Zeng et al., 2017).

Related Work

A system associated with water and its interconnections is called hydrological system. Analysis and modeling of hydrological systems have various applications in many fields such as (Thomann, 1967), (Lohani and Wang, 1987), and (Huck and Farquhar, 1974) in water quality modelling; (Srikanthan and McMahon, 1982) in annual and monthly rainfall simulation; (McMicheal and Hunter, 1972) in modeling temperature and flows in rivers. In the context of Pakistan, monsoon precipitation plays a significant part in water flow in Indus river basin from April to September each year. Yearly rainfall varies from 1000-1400 mm in the entire regions. This needs a sufficient model development of stream flow on the Indus basin for estimating and enhancing the stream structures in Pakistan so that it may help to water resource operators and planners.

There are different methods for estimation of hydrological analysis, such as bivariate frequency analysis (Yen and Rasmussen, 2002). A study computed the possibility of flood in various dams on Indus River basin in Pakistan (Khan, Iqbal and Yosufzai, 2011). A research has been developed for nonlinear forecasting techniques to examine the stream flow on the Indus River basin (Hassan and Ansari, 2010). On the other hand, (Sudheer, Nayak and Rangan, 2000) have accomplished a deep analysis on the flood risk for the basin by employing a simulated artificial neural system (ANN) model and they described that there is a space for more improvement in ANN model.

Data Description and Locations

Pakistan has natural diversity of climatic conditions in different parts of the country and observed the extreme variations in hydro-climatic variables. The climate extremes in Pakistan incorporate high and low temperatures, heaviest precipitation or rainfall etc. which causes floods that are regularly activated by strong Indus River basin Fig. 3.1 which is an important trans-limit stream in Asia with nine tributaries.

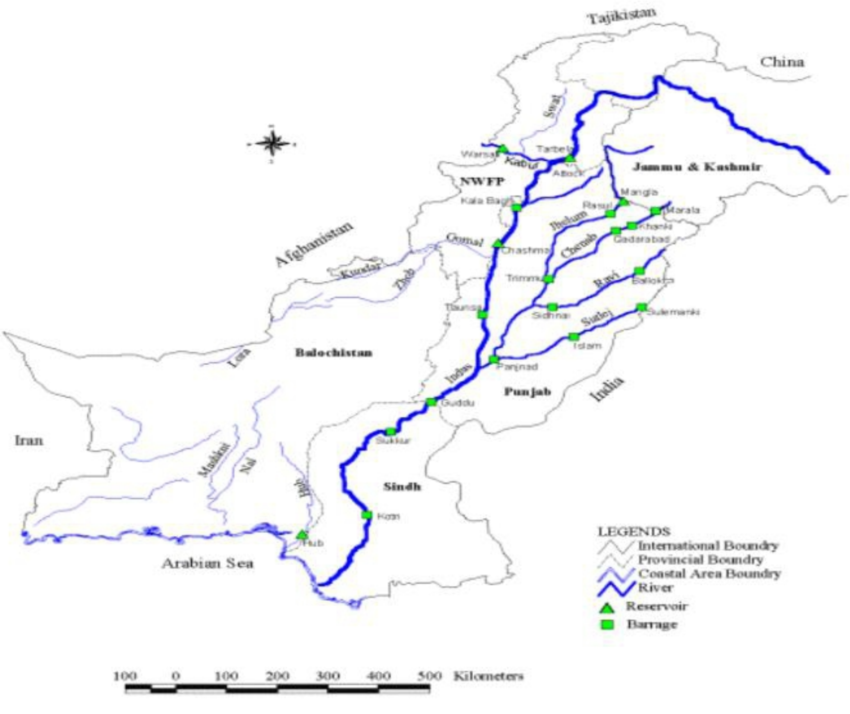

Figure 1: Indus River basin in Pakistan

This paper considers mean monthly upstream flow observations for the gauge station at Sukkur barrage on lower Indus river basin from April to

September (due to pre-monsoon and monsoon rainfall) each year, ranging from 1977 to 2012.

Dynamic Neural Network

Based on the learning aspect of human computational abilities, Dynamic Neural Network (DNN) is a type of Artificial Neural Network (ANN) in the paradigm of computational intelligence. The neuron is the fundamental structure component of the ANN. The following Fig. 4.1 exhibits the basic neuron model which consist of the inputs $I_1, I_2, \ldots, I_n$ and the output O. The body adds the weighted inputs and evaluates the sum against the threshold value (θ).

That is the total weighted input = $I_1 w_1 + I_2 w_2 + \ldots\ldots\ldots\ldots.. + I_n w_n$

Total weighted input = $\sum_{i=1}^{n} I_i w_i$

(4.1)

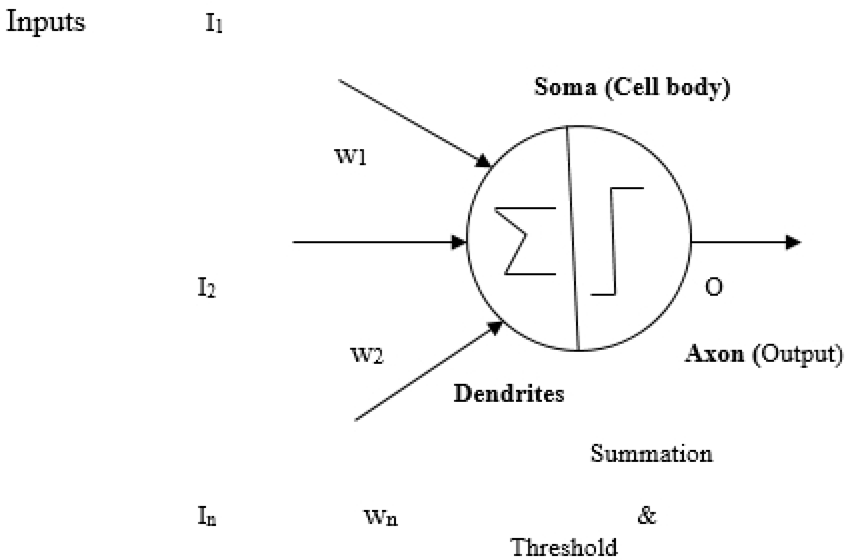

Figure 2: Fundamental neuron model

The output is showed by:

$$O = f_h\left(\sum_{i=1}^{n}(I_i w_i - \theta)\right)$$

(4.2)

where f_h is a Heaviside function, expressed by:

$$f_h = \begin{cases} 1, & if\ x > 0 \\ 0, & if\ x \leq 0 \end{cases}$$

(4.3)

The bias θ is expressed as an input with weight equal to 1, so equation (2.2) becomes:

$$O = f_h\left(\sum_{i=1}^{n}(\,I_i w_i)\right)$$

(4.4)

This neuron model put forward by (McCulloch and Pitts, 1943). The other transfer functions are also used which include linear and sigmoid functions. The sigmoid function is expressed as:

$$y = \frac{1}{1 + e^{-kI}}$$

(4.5)

Where y is the output, I is the total input and k is a gain factor that controls the sharpness of the transition from 1 to 0. The sigmoid function is principally of use because it is easy to compute the derivative so that it may save computational time execution for training algorithms.

$$\text{If } y = \frac{1}{1+e^{-kI}}, \text{ then } \frac{dy}{dI} = ky(1-y)$$

(4.6)

A good training algorithm of the feed-forward network is generally Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963) because it is fastest among all the training algorithms.

Recommended Models Based on DNN

This paper, suggests three Dynamic Neural Network (DNN) models Table 1, based on time delayed autoregressive structures for upstream water flow at Sukkur barrage on lower Indus River basin in Pakistan. The appropriateness of the models for training, validation and testing phases established on evaluation metrics which exhibit the efficiency and accuracy of the models.

Table 1: Dynamic Neural Network (DNN) functional models

Model No.	Delayed Inputs Functional Models
1	$Q_n = f(\, Q_{n-2}, Q_{n-1}\,)$
2	$Q_n = f(Q_{n-3}, Q_{t-2}, Q_{t-1})$
3	$Q_n = f(Q_{n-4}, Q_{n-3}, Q_{n-2}, Q_{n-1})$

Where

Q_n = Mean monthly upstream water flow at time n

Q_{n-1} = Mean monthly upstream water flow at time n-1(lag 1)

Q_{n-2}= Mean monthly upstream water flow at time n-2 (lag 2)

Q_{n-3}= Mean monthly upstream water flow at time n-3 (lag 3)

Q_{n-4}= Mean monthly upstream water flow at time n-4 (lag 4)

Training Factors:

Number of epochs	1000
Minimum gradient	1.0e-07
Initial step size	0.001
Step incrementing rate	10
Step decrementing rate	0.10
Maximum step size	1.0e+10

Training Function:
Levenberg-Marquardt

Results and Discussion

The numerical results of four dynamic neural networks (DNN) models based on variable parameters (number of neurons in hidden layer and time delay in months) for upstream flow at Sukkur Barrage on Indus River basin are shown in the forms of figures and tables. The models evaluation metrics in Tables 2 and 3 exhibit that model 1 is a good agreement between forecasted and observed data as compare to other models.

DNN Model 1

Hidden layer (neurons)	14
Time delay (months)	2

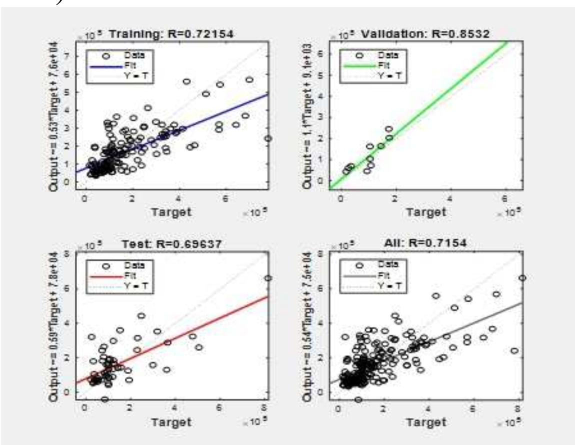

Figure 3: Output of DNN model 1 of Sukkur Barrage

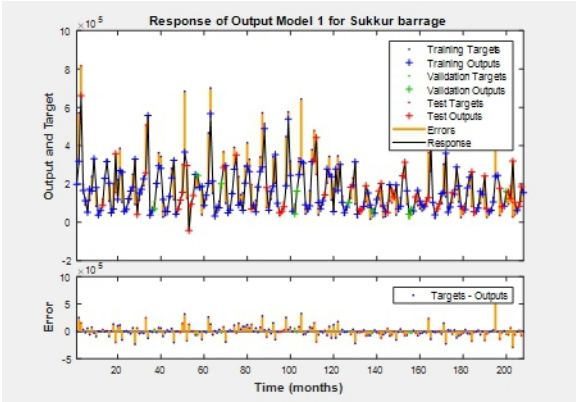

Figure 4: Response and errors of output model 1 of Sukkur Barrage

DNN Model 2

Hidden layer (neurons)	11
Time delay (months)	03

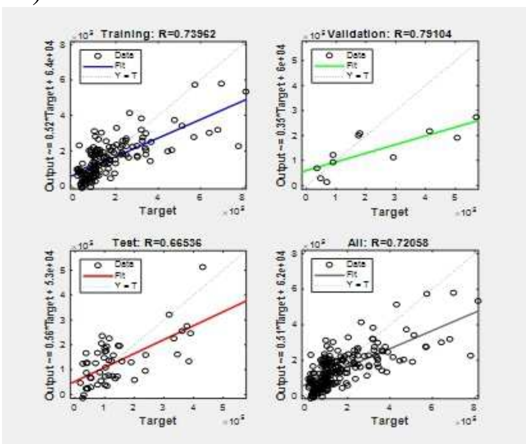

Figure 5: Output of DNN model 2 of Sukkur Barrage

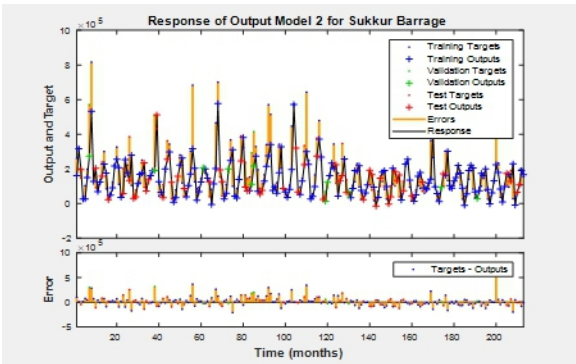

Figure 6: Response and errors of output model 2 of Sukkur Barrage

DNN Model 3

Hidden layer (neurons) 13

Time delay (months) 4

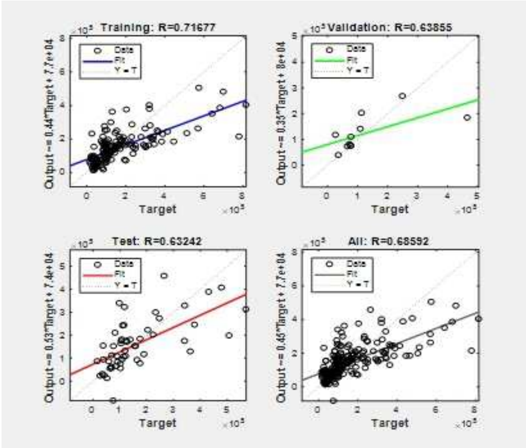

Figure 7: Output of DNN model 3 of Sukkur Barrage

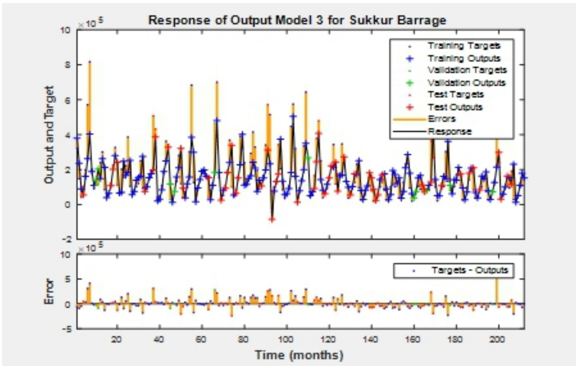

Figure 8: Response and errors of output model 3 of Sukkur Barrage

<u>Models Evaluation Metrics for Sukkur Barrage</u>

The following two tables of evaluation metrics of DNN models are given for mean monthly upstream flow at Sukkur Barrage and will select a network model on the basis of the following criteria:

a. Network model has maximum value of the coefficient of correlation (R) on the testing or forecasting phase.
b. Network model has maximum value of the coefficient of determination (R^2) on the testing or forecasting phase.
c. Network model has minimum error (RMSE) on the testing or forecasting phase.

Table 2: Coefficient of Correlation (R) and Coefficient of Determination (R^2) Evaluation Metrics of All DNN Models of Sukkur Barrage

Model No.	**Training (R)**	**Validation (R)**	**Testing (R)**	**Training (R^2)**	**Validation (R^2)**	**Testing (R^2)**
1	0.7215	0.8532	**0.6964**	0.5205	0.7279	**0.4849**
2	0.7396	0.791	0.6653	0.547	0.6256	0.4426
3	0.7167	0.6385	0.6324	0.5136	0.4076	0.399

Table 3: RMSE Evaluation Metric of All DNN Models of Sukkur Barrage

Model No.	**Training (RMSE)**	**Validation (RMSE)**	**Testing (RMSE)**
1	46340.90	39221.92	**46340.95**
2	104749.39	156283.60	84302.59
3	110379.27	94151.39	103087.06

Conclusion

A standard method of computational or artificial intelligence, called Dynamic Neural Network (DNN), has been developed to handle the issues of nonlinearity and nonstationarity in the real world complex system. This paper Presents the pursued of our research for developing Neural Network based models which could predict or forecast mean monthly upstream flow through delayed or lagged variables. The nonlinear autoregressive (NAR) based four different DNN models are inferred or extracted with the optimized parameter of number of neurons in a hidden layer. By using the three measures of goodness of fit through models' evaluation metrics over training, validation, and testing sets, it is concluded that the values of the performance measures of coefficient of correlation and determination are sufficient or adequate to obtain the DNN models in which model one is more efficient as an appropriate autoregressive structure based dynamic neural network model for mean monthly upstream flow at Sukkur barrage on Indus river basin and it may be more useful for water resource management of the basin.

References:

1. Ben Taieb, S., Bontempi, G., Atiya, A. F., and Sorjamaa, A. (2012). A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. Expert Systems with Applications, 39(8), 7067–7083.
2. Bodyanskiy, Y., and Popov, S. (2006). Neural network approach to forecasting of quasi periodic financial time series. European Journal of Operational Research 175(3), 1357–1366.
3. Chen, C. F., Lai, M. C., and Yeh, C. C. (2012). Forecasting tourism demand based on empirical mode decomposition and neural network. Knowledge-Based Systems, 26, 281–287.

4. Cherif, A., Cardot, H., and Boné, R. (2011). SOM time series clustering and prediction with recurrent neural networks. Neurocomputing, 74 (11), 1936–1944.
5. Hassan, S. A., and Ansari, M. R. K. (2010). Nonlinear analysis of seasonality and stochasticity of the Indus River. Hydrological Sciences Journal, 55(2), 250-265.
6. Hu, Y. C. (2017). Electricity consumption prediction using a neural-network-based grey forecasting approach. Journal of the Operational Research Society, 68(10), 1259–1264.
7. Huck, P. M., and Farquhar, G. J. (1974). Water quality models using Box-Jenkins method. Journal of Environmental Engineering, ASCE, 100(3), 733–753.
8. Kantelhardt, J. W., Zschiegner, S. A., Koscielny-Bunde, E., Havlin, S., Bunde, A., and Stanley, H. E. (2002). Multifractal detrended fluctuation analysis of nonstationary time series. Physica A, 316, 87–114.
9. Khan, B., Iqbal, M. J., and Yosufzai, M. A. K. (2011). Flood risk assessment of river Indus of Pakistan. Arab. J. Geosci., 4(1-2), 115-122.
10. Lai, Y., Ye, N. (2003). Recent developments in chaotic time series analysis. International Journal of Bifurcation & Chaos, 13(06) 1383–1422.
11. Levenberg, K. (1944). A Method for the Solution of Certain Non-Linear Problems in Least Squares. Quarterly of Applied Mathematics, 2, 164–168.
12. Li, G., and Song, H. (2008). Tourism demand modeling and forecasting-A review of recent research. Tourism Management, 29(2), 203–220.
13. Lohani, B. N., and Wang, M. M. (1987). Water quality data analysis in Chung Kang River. Journal of Environmental Engineering, ASCE, 113(1), 186–195.
14. Luger, G. F. (2005). Artificial Intelligence: Structures and Strategies for Complex Problem Solving. Fifth edition, Addison-Wesley.
15. Lu, X., Wang, J., Cai, Y., and Zhao, J. (2015). Distributed HS-ARTMAP and its forecasting model for electricity load. Applied Soft Computing, 32, 13–22.
16. Marquardt, D. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. SIAM Journal on Applied Mathematics, 11(2), 431–441. doi:10.1137/0111030.
17. McCulloch, W. S., and Pitts, W. H. (1943). A logical calculus of ideas immanent in nervous activity. Bulletin of Mathematical Biophysics, 5, 115–133.

18. McMicheal, F. C., and Hunter, J. S. (1972). Stochastic modeling of temperature and flow in rivers. Water Resources Research, 8(1), 87–98.
19. Srikanthan, R., and McMahon, T. A. (1982). Simulation of Annual and Monthly Rainfalls—A preliminary study at five Australian stations. Journal of Applied Meteorology, 21, 1472–1479.
20. Sudheer, K. P., Nayak, P. C., and Rangan, M. (2000). Rainfall runoff modeling using artificial neural network technique. Report No. CS/AR-16/1999-2000, National Institute of Hydrology, Roorkee, India.
21. Thomann, R.V. (1967). Time series analysis of water quality data. Journal of Sanitary Engineering Division, ASCE, 93(1), 1–23.
22. Wan, Y., Gong, X., and Si, Y. (2016). Effect of segmentation on financial time series pattern matching. Applied Soft Computing, 38, 346–359.
23. Wang, L., Wang, Z., and Liu, S. (2016). An effective multivariate time series classification approach using echo state network and adaptive differential evolution algorithm. Expert Systems with Applications, 43, 237–249.
24. Yen, S., and Rasmussen, P. (2002). Bivariate frequency analysis: discussion of some useful concepts in hydrological application. Hydrological Processes, 16(14), 2881-2898.
25. Zeng, Y. R., Zeng, Y., Choi, B., and Wang, L. (2017). Multifactor-influenced energy consumption forecasting using enhanced back-propagation neural network. Energy, 127, 381–396.
26. Zhang, J., Luo, X., Nakamura, T., Sun, J., and Small, M. (2007). Detecting temporal and spatial correlations in pseudo periodic time series. Physical Review E Statistical Nonlinear & Soft Matter Physics. 75(2), 016218.