ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 5, Issue 9 , September 2018

Daily Streamflow Forecasting Using Nonlinear Echo State Network

S. Bahrami*, E. P. Wigand

Department of Hydrological Science, College of Science, University of Nevada, Reno
Department of Geography, College of Science, University of Nevada, Reno

ABSTRACT: The prediction of streamflow is an important issue in hydrologic engineering and hydropower reservoirmanagement.Several approaches including statistical, physical or conceptual models have been investigated to forecast streamflow. Most of the methods assume a linear relationship between the input and output series. However, they ignore the nonlinear information hidden in the streamflow series .In this paper, various time series inputs including: day length, precipitation, solar radiation, maximum and minimum temperature per day, and vapor pressure have been used. An advanced and powerful forecast engine called Nonlinear Echo State Network using Multivariable Polynomial (NESN-MP)is used to predict the behaviour of the streamflow. The forecasting is conducted under different climatic conditions to indicate the model's applicability. Furthermore, to demonstrate the efficiency of the proposed method, it is compared with Adaptive Neuro-Fuzzy Inference System (ANFIS). The results of the new method compare favourably with ANFIS.

KEYWORDS: Forecasting, nonlinear echo state network using multivariable polynomial (NESN-MP), streamflow.

I. INTRODUCTION

State estimation and forecasting of streamflow have always been general concerns for engineers. State estimation is applied in all energy management systems to identify the present operating state of a system [1-2]. Forecasting is also an important and necessary aid to planning and planning is the backbone of effective operations. In hydrology, streamflow forecasting is vital for water resources engineers, reservoir operators and water managers who strive to balance a range of competing objectives to support their decisions about hydroelectric power programming, flood mitigation, agricultural and domestic water supplies, irrigation management as well as maintenance of environmental flows [3]. Therefore, developing an optimal streamflow forecasting model as a stochastic property of environmental modelling is crucial. However, the existing dynamicity, inherent complexities and chaotic feature in the temporal and spatial expansion of the model may obstruct the accurate and reliable prediction process [4].

Different statistical, physical or conceptual models have been evolved to forecast streamflow [5]. Statistical models, such as regression-based models [6]are extremely simplistic and suffer from a functional form between variables prior to the analysis. Therefore, they do not properly account for the relationship between the dependent and observed explanatory variables. Physically based numerical models, typically, simulate the streamflow generation process through a governing equation employing limited boundary conditions, which need precise data input to enable parameter calibration[7].Conceptual hydrological models consider different processes of the hydrological cycle along with mathematical formulation to improve the forecasting accuracy[8]such as:the Soil and Water Assessment Tool (SWAT) as a semi-distributed conceptual model [9]. Louise J. et al 2017 [10] conducted research to forecast streamflow in deterministic and probabilistic terms for all initialization months, flow quintiles, and seasons. The result showed a relatively accurate streamflow forecasts from low to high flows, but their model could not decrease uniformly with initialization time.

However, taken as a group, the accuracy of these models is not reliable due to heterogeneous hydrogeological characteristics within the watershed system in nature with respect to time and space. In addition, large data input, large number of parameters, and broad range of necessary values may limit the application of comprehensive simulation models [11]. Furthermore, all of these models assume that the relationship between the input and output series is linear or at worst near linear. They thus ignore the nonlinear information hidden in the streamflow series which result in a poor model performance. Furthermore, streamflow is under the influence of many factors such as evapotranspiration, rainfall, atmospheric circulation and temperature, and its generation process is nonlinear and time-variable. Therefore,

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 5, Issue 9 , September 2018

in the last two decades, researchers have focused on alternative data-base predictive methods. Several studies have been presented on developing soft computing tools with Artificial Intelligence (AI) models. Several computer models have been recently conducted to forecast streamflow and corresponding runoff. Some of these models are based on the Artificial Neural Networks (ANN), fuzzy network sets, genetic programming, regression algorithms, support vector machine; and nonparametric methods such as K-Nearest Neighbour (KNN) Regression [12-13]. Among all mentioned methods, the fuzzy network sets depend on the user expertise, while the efficiency of others depends on the model ability to find out the relationship between input and output variables. It has been demonstrated that Support Vector Machine (SVM) yields the most accurate results compared to Auto Regressive Moving Average (ARMA), ANN, and Multiple Linear Regression (MLR) [14], and (something missing?)[12]. On the other hand, Shrestha (2014) indicated that the result of annual predicted streamflow using ANN and SVM, throughout the State of Utah, are the same [15]. Yong Liu et.al (2016) compared the RVM and SVM for long term streamflow forecasting. They found that RVM produces better results for annual streamflow forecasting within a specified climatic condition [16]. Bharti et al. (2017) indicated that in forecasting process of monthly runoff, ANN results surpasses the Least Square- Support Vector Regression (LS-SVR) results, while LS-SVR results exceed ANN results for monthly sediment prediction [17].

ANNs are the most popular artificial intelligence (AI) techniques used in variety of fields especially in time series forecasting. Successful prediction results of ANN application in hydrological process such as rainfall-runoff modelling, streamflow prediction, reservoir inflow forecasting, rainfall forecasting, and river sediment modelling have been recently published. Although different feed forward neural network models have been well documented, the selection basis of these models has thus farreceived limited attention [18-19]. Kerh and Lee (2006) introduced ANNs to predict flood discharge at downstream stations with data scarcity, using information at upstream stations of the Kaoping River [20]. Their model demonstrated that back-propagation of the ANN model performs better than the conventional Muskingum method. Due to chaotic behaviour in hydrological time series, one of the most important steps in constructing an ANN model for streamflow forecasting is determining the best inputs. Zhao, X. at.al (2017) used the Phase Space Reconstruction (PSR) method as an alternative approach to select relevant and important input variables for ANN models. They built two different ANN models using the time-lagged records of precipitation and temperature. They indicated that ANNs predict daily streamflow in the adjacent ungauged basins as accurate as in the gauged basin [21]. Zealand et al (1999) used the ANN trained with back-propagation algorithm to predict streamflow 1-week-ahead [22]. However, ANN models have some lapses including over-fitting and under-fitting, slow learning speed, and curse of dimensionality and convergence to local minimum. Therefore, in processing of complex hydrological phenomena, they betraya poor performance [23-25].Typically, their disadvantages include the following [26]:

- *High complexity and long processing time.*
- *High dependence on parameter tuning and optimization.*
- *The requirement for nonconvex optimization that can yield suboptimal results and trap in local optima.*

In this paper, NESN-MPhas been used as a forecasting engine. The network consists of a reservoir including linear internal states and a readout including nonlinear functions of the internal state. The nonlinear relations between the internal states increase the learning capability, which results in high forecasting accuracy while ensuring that the quality of forecasting does not deteriorate significantly with time. Furthermore, the performance of the forecasting engine is improved by decreasing the number of internal states, and the orders of the weight matrices, which reduces the computational load considerably. Furthermore, in all previous research, the results obtained from these studies are inconsistent due to difference in study areas, input data sets, and the selected structures for each of the models [27]. Many studies have applied the original streamflow time series as the input variables in their forecasting model, which results in missing some features of different resolution [28]. Using just one resolution component could not reflect the internal mechanism of streamflow .Therefore, daily data is preferred, because it is not significantly affected by external factors such as meteorological pattern and anthropogenic activities in the data [29]. However, research has been conducted on evaluation of annual, seasonal or monthly streamflow for one-time scale condition [25].Moreover, daily streamflow forecasting at different time scales has not been addressed in the literature.

Based on the outline above, the study in this paper has developed a modified model of ANN to forecast daily streamflow based on various time series forcing-data input including daily precipitation, precipitation duration, solar radiation, temperature and vapor pressure. The remainder of this study is as follows. Section II provides an overview of the NESN. Simulation results are given in Section III, and conclusions are summarized in Section IV.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 5, Issue 9 , September 2018

II. NONLINEAR ECHO STATE NETWORK

This powerful method is simple, effective, with far fewer computations [30]. NESN-MP provides a total of $2p + p^2$ units;$p$internal states; p squares of the internal states; and p^2 units gained by multiplying the internal states and squares of the internal states. Therefore, the order of weight matrices is decreased radically. The weight matrices (W, T, and V) are then used to calculate the internal states of the reservoir. The vector of internal states is updated using

$$\boldsymbol{x}_{(t+1)} = f(W.\boldsymbol{x}_{(t)} + V.\boldsymbol{s}_{(t+1)} + T.\boldsymbol{y}_{(t)}) \tag{1}$$

and the readout vector is

$$\overline{\boldsymbol{x}}_{(t+1)} = [\boldsymbol{x}_{(t+1)} \quad , \boldsymbol{x}^2{}_{(t+1)} \quad , \sum_{i_1=1}^{p} \sum_{i_2=1}^{p} \boldsymbol{x}_{i_1 (t+1)} . \boldsymbol{x}^2{}_{i_2 (t+1)}] \tag{2}$$

where $\boldsymbol{x}^2{}_{(t+1)} = \left[x^2_{1(t+1)}, x^2_{2(t+1)}, \dots, x^2_{p(t+1)}\right]$, p is the number of internal states $\left[\frac{N}{p+2}\right]$, $\boldsymbol{s} \in R^{K\times 1}$ is the input vector,$\boldsymbol{x} \in R^{p\times 1}$ is the internal state vector, $\overline{\boldsymbol{x}} \in R^{(p^2+2p)\times 1}$ is the readout vector, and$\boldsymbol{y} \in R^{L\times 1}$ denotes the output states.

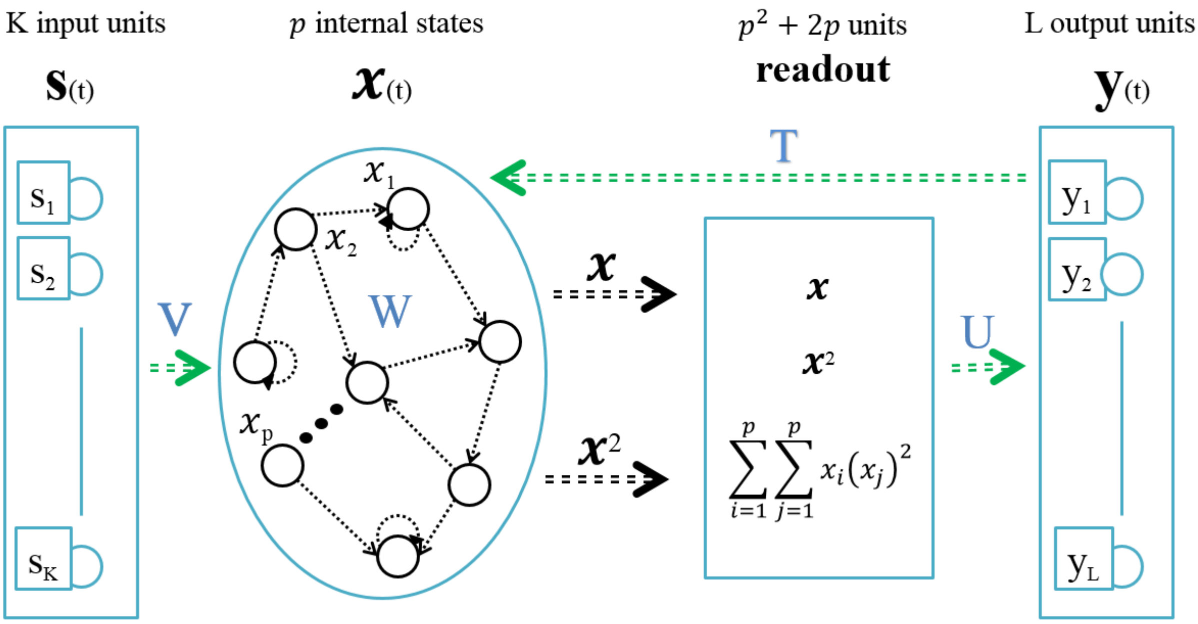

Fig. 1. Schematic of NESN-MP.

The matrix $W \in R^{p\times p}$ defines the internal state interconnections within the reservoir. The values in W are fixed values generated randomly over a symmetric interval.

$$W = (w_{ij})_{p\times p} \;\; ; \; w_{ij} \in (-1,1)(i, j = 1,2, \dots, p) \tag{3}$$

Matrix $V \in R^{p\times K}$, containing randomly chosen fixed values, defines the connections of the input with the internal states of the reservoir.

$$V = (v_{ij})_{p\times k} \;\; ; v_{ij} \in (-1,1)(i = 1,2, \dots, p, j = 1,2, \dots, k) \tag{4}$$

The output feedback matrix,$T \in R^{p\times L}$ is

$$T = (t_{ij})_{p\times L} \;\; ; \; t_{ij} \in (-1,1)(i = 1,2, \dots, p, j = 1,2, \dots, L) \tag{5}$$

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 5, Issue 9 , September 2018

The output matrix, $U \in R^{L\times(p^2+2p)}$ is

$$U \quad = (u_{ij})_{L\times(p^2+2p)} \; ; u_{ij} \in (-1,1)(i = 1,2,\dots,L, j = 1,2,\dots,2p + p^2) \tag{6}$$

where K is the number of inputs, p is the number of internal states, and L is the number of outputs.

III. SIMULATION RESULTS

The performance of the NESN-MPis tested using climatic observation data (day length, precipitation, solar radiation, maximum and minimum temperature per day, and vapor pressure) with a time interval of 24 hours used to train and test the proposed methods. Each data set is divided into two separate parts for training and testing, with their lengths denoted as L_{train} and L_{test}, respectively. The MSE, root mean squared error (RMSE), normalized root-mean-square error (NRMSE), normalized mean-absolute error (NMAE), and mean absolute error (MAE) shown in (7-11) were used to evaluate the performance of the proposed methods.

NRMSE is often expressed as a percentage and calculated as

$$\text{NRMSE} = \sqrt{\frac{\sum_{i=1}^{n_{max}} ||y(i) - \hat{y}(i)||^2}{\sum_{i=1}^{n_{max}} ||y(i) - \tilde{y}||^2}} \times 100\% \tag{7}$$

where $\|\bullet\|$ indicates the Euclidean norm, $\boldsymbol{y}$ are the actual output values, $\tilde{y}$ is the average of $\boldsymbol{y}$ over the whole target set $y(1), y(2), \dots, y(n_{max}), \hat{\boldsymbol{y}}$ is the predicted output, and $n_{\max}$ is the number of sample points. Lower values in NRMSE indicate less residual variance. In many cases, especially for smaller samples, the sample range is likely to be affected by the size of sample, which would hamper comparisons.

MSE measures the average of the squares of the errors, which is always non-negative, and values closer to zero are better. Taking the square root of MSE yields RMSE, which has the same units as the estimated quantity. MSE and RMSE are calculated as

$$\text{MSE} = \frac{1}{n_{max}} \sum_{i=1}^{n_{max}} [y(i) - \hat{y}(i)]^2 \tag{8}$$

$$\text{RMSE} = \frac{1}{n_{max}} \sqrt{\sum_{i=1}^{n_{max}} [y(i) - \hat{y}(i)]^2} \tag{9}$$

MAE calculates the average magnitude of the errors in a set of predictions without considering their direction. It is the average over the test sample of the absolute differences between prediction and actual observation where all individual differences have equal weight. NMAE normalizes MAE by the range of available rating values. MAE and NMAE are defined as

$$\text{MAE} = \frac{1}{n_{max}} \sum_{i=1}^{n_{max}} |y(i) - \hat{y}(i)| \tag{10}$$

$$\text{NMAE} = \frac{1}{y_{max} \cdot n_{max}} \sum_{i=1}^{n_{max}} |y(i) - \hat{y}(i)| \tag{11}$$

where $y_{\max}$ is the maximum value of output. Generally, RMSE and MAE are regularly employed in model evaluation studies [26].

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 5, Issue 9 , September 2018

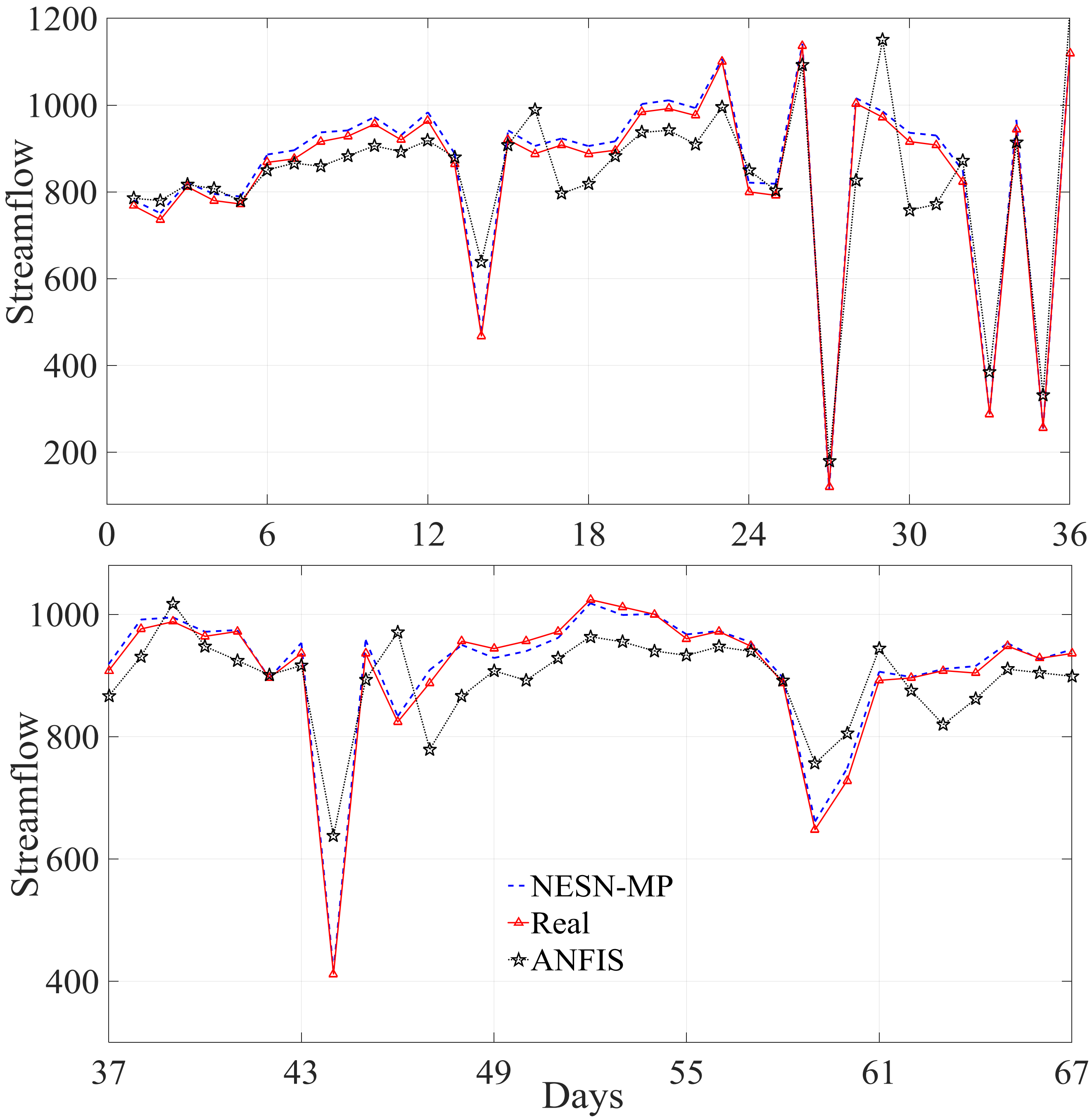

Fig 2. 67 days prediction in case study in 1995.

The streamflow forecasting is carried out for 67 days ahead. $l_{train} = 200$, $l_{test} = 67$ with no overlap and with the test data starting immediately after the training data. Fig. 2shows the prediction for 67 days ahead for NESN-MP and ANFIS. NESN-MP provide an MAE of 4 for the first 10 days ahead which is significantly below the MAE given by ANFIS. This improvement can be seen on the second 10 days more obviously, where MSE and MAE in ANFIS results increased considerably while those in the proposed methods remain almost constant. The results clearly show that the proposed NESN-MP outperform ANFIS.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 5, Issue 9 , September 2018

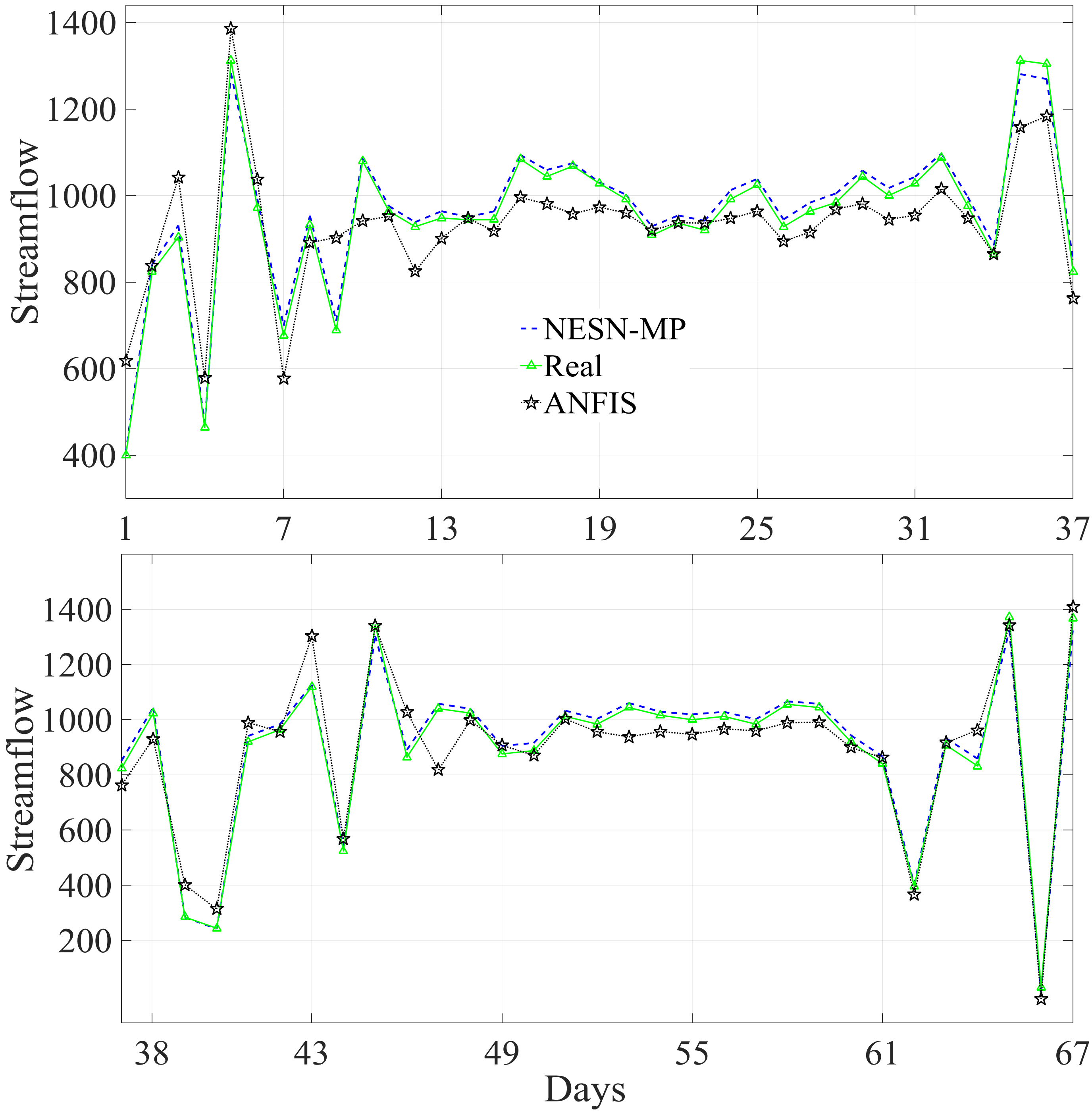

Fig 3. 67 days prediction in case study in 2005.

To validate the prediction ability and universality of the proposed methods with different climatic parameters, 67 days streamflow forecasting in 2005 is shown in Fig. 3. It is shown that the NESN-MP provide an MAE of 9 for the first 10 days ahead forecasting which is 92.8% below the MAE for ANFIS, respectively. In case of RMSE, NESN-MP gives the respective value of 11.21 which are well below the RMSE of 101.2 for ANFIS. Table 1 shows the error indices for both methods for different days.

ISSN: 2350-0328
International Journal of AdvancedResearch in Science, Engineering and Technology
Vol. 5, Issue 9 , September 2018

Table 1. Error indices for case study 2005.

		Days 1-10	Days 11-20	Days 21-30	Days31-40	Days 41-50	Days 51-60
ANFIS	**MSE**	10235	3437	1665	7600	2347	763
	RMSE	101.2	58.62	40.8	87.18	48.45	27.63
	NMAE%	16.22	10.98	7.84	15.93	8.8	5.55
	NRMSE %	9.2	5.329	3.7	7.92	4.4	2.51
	MAE	73	49.4	35.3	71.7	39.6	25
NESN-MP	**MSE**	125	43.2	219	264	94.5	80.8
	RMSE	11.21	6.57	14.8	16.26	9.7	8.98
	NMAE%	2	1.2	2.95	2.7	1.89	1.82
	NRMSE %	1.02	0.6	1.34	1.47	0.88	0.81
	MAE	9	5.4	13.3	12.2	8.5	8.2

IV. CONCLUSION

This study presents daily streamflow forecast based on various time series forcing-data inputs including daily precipitation, precipitation duration, solar radiation, temperature and vapor pressure. A novel echo state networks called NESN-MP has been used as forecasting engine. The nonlinear relations between the internal states increase the learning capability, which results in high forecasting accuracy while ensuring that the quality of forecasting does not deteriorate significantly with time. Furthermore, the daily values for different parameters which affect the streamflow provide accurate forecasting. Simulation results validate the performance of the proposed method and demonstrate its superiority over ANFIS. NESN-MP provides significantly lower values than those given by ANFIS for MAE, NMAE, MSE, RMSE, and NRMSE. Future work will compare the proposed method with the classical methods.

REFERENCES

[1] M. A. Chitsazan, M. S. Fadali, A. M Trzynadlowski, “State estimation of IEEE 14 bus with unified interphase power controller (UIPC) using WLS method”, Energy Conversion Congress and Exposition (ECCE), 2017 IEEE, pp. 2903-2908, Oct. 2017.
[2] M. A. Chitsazan, A. M Trzynadlowski, “State estimation of IEEE 14 bus with interphase power controller using WLS method”, Energy Conversion Congress and Exposition (ECCE), 2016 IEEE, pp. 1-5, Sep. 2016.
[3] Dehghani, M., Saghafian, B., Nasiri, S. F., Farokhnia, A., and Noori, R. “Uncertainty analysis of streamflow drought forecast using artificial neural networks and Monte-Carlo simulation.”, Int. J. Climatol,2014,pg no : 1169–1180
[4] Bayazit, M.” Nonstationarity of Hydrological Records and Recent Trends in Trend Analysis: A State-of-the-art Review”, Environ. Process, 2015, pg no : 527–542.
[5] Partington, D., Brunner, P., Simmons, C.T., Werner, A.D., Therrien, R., Maier, H.R., Dandy, G.C. “Evaluation of outputs from automated baseflow separation methods against simulated baseflow from a physically based, surface water groundwater flow model”, Journal of Hydrology, 2012, pg no : 28-39.
[6] Chua, I.H.C., Wong, T.S.W.”Runoff forecasting for an asphalt plane by artificial neural networks and comparisons with kinematic wave and autoregressive moving average models”, Journal of Hydrology, 2011, pg no : 191–201
[7] Yoon, H., Jun, S.C., Hyun, Y., Bae, G.O., Lee, K.K. “A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer”,Journal of Hydrology, 2011, pg no : 128–138.
[8] Peng, T.; Zhou, J.; Zhang, C.; Fu, W. “Streamflow Forecasting Using Empirical Wavelet Transform and Artificial Neural Networks”. Water, 2017.
[9] Arnold, J.G., R. Srinivasan, R.S. Muttiah, and J.R. Williams “ Large Area Hydrologic Modeling and Assessment. Part I: Model Development”, Journal of the American Water Resources Association (JAWRA) ,1998, pg no :73-89
[10] Strayer, D. L., M. E. Power, W. F. Fagan, and S. T. A. Pickett. “A classification of ecological boundaries”, BioScience ,2003, pg no:723–729.
[11] Bhattacharya, B., Solomatine, D.P. “Application of artificial neural network in stage–discharge relationship”, In: Proceedings of the 4th International Conference on Hydro-informatics, Iowa City, IA, 2000, pg no : 1–7
[12] Ahmed, J.A., Sarma, A.K.” Artificial neural network model for synthetic streamflow generation”,Journal of Water Resources Management, 2007, pg no : 1015-1029.
[13] Kisi, O. “Streamflow forecasting and estimation using least square support vector regression and adaptive neuro-fuzzy embedded fuzzy c-means clustering”, Journal of Water Resources Management, 2015, pg no : 5109–5127
[14] Kalra, A., W. P. Miller, K. W. Lamb, S. Ahmad, and T. Piechota.“Using large-scale climatic patterns for improving long lead time streamflow forecasts for Gunnison and San Juan River Basins”, Hydrol. Processes, 2013, pg no : 1543–1559
[15] Shrestha, A. B. & Bierkens, Lutz, A. F., Immerzeel, W. W., M. F. P. “Consistent increase in High Asia’s runoff due to increasing glacier melt and precipitation”, Nat. Clim. Chang, 2014, pg no : 587–592
[16] H Zhang, R Banerjea, Y Liu, M Yu.” Number of streams indication for WLAN”, US Patent 9,294,164, 2016.
[17] Bharti B, Goyal MK, Quilty J, Adamowski J, Pandey A. “Modeling of daily pan evaporation in sub-tropical climates using ANN, LS-SVR, Fuzzy Logic, and ANFIS”, Expert Syst Appl ,2014, pg no :5267–5276

ISSN: 2350-0328

International Journal of AdvancedResearch in Science, Engineering and Technology

Vol. 5, Issue 9 , September 2018

[18] Shoaib M, Shamseldin AY, Khan S, Khan MM, Khan ZM, Sultan T, Melville BW. "A comparative study of various hybrid wavelet feedforward neural network models for runoff forecasting", Water Res Manage,2018, pg no : 83–103

[19] Imrie, C. E., Durucan, S. & Korre, A. "River flow prediction using artificial neural network generalisation beyond calibration range",Journal of Hydrology, 2000, pg no : 138–153

[20] Kerh T, Lee CS. "Neural networks forecasting of flood discharge at an unmeasured station using river upstream information", Adv Eng Softw, 2006, pg no :533–543.

[21] Zhao Y, Wei Y, Li S, Wu B."Downstream ecosystem responses to middle reach regulation of river discharge in the Heihe River Basin, China", Hydrol Earth Syst Sci, 2016, pg no :4469-4481

[22] Zealand, C.M., Burn, D.H., Simonovic, S.P."Short term streamflow forecasting using artificial neural networks", Journal of Hydrology, 1999, pg no : 32–48.

[23] Huang, S., Chang, J., Huang, Q., Chen, Y. "Monthly streamflow prediction using modified EMD-based support vector machine".. Journal of Hydrology,2014, pg no : 764–775.

[24] Nourani, V., Baghanam, A., Adamowski, J., Kisi, O.,"Applications of hybrid wavelet-artificial intelligence models in hydrology: A review", Journal of Hydrology ,2014, pg no : 358–377

[25] Niu, W.J., Feng, Z.K., Cheng, C.T., and Zhou, J.Z. "Forecasting daily runoff by extreme learning machine based on quantum-behaved particle swarm optimization.",J. Hydrol. Eng , 2018, pg no : 04018002-10

[26] M. A. Chitsazan, M. Sami Fadali, A. M Trzynadlowski, "Wind speed and wind direction forecasting using echo state network with nonlinear functions", Renewable Energy, 2019, pg no :879-889.

[27] Modaresi, F.; Araghinejad, S.; Ebrahimi, K. "A Comparative Assessment of Artificial Neural Network, Generalized Regression Neural Network, Least-square Support Vector Regression, and K-nearest Neighbor Regression for Monthly Streamflow Forecasting in Linear and Nonlinear Conditions", Water Resource Manage, 2018, pg no : 243–258

[28] Guo, X., Hu, T., Wu, C., Zhang, T., & Lv, Y. "Multi-objective optimization of the proposed multi-reservoir operating policy using improved NSPSO", Water Resources Management, 2013,pg no : 2137–2153.

[29] Yaseen, Z.M.; Fu, M.; Wang, C.; Hanna, W.; Wan, M.; Deo, R.C.; El-shafie, A. "Application of the Hybrid Artificial Neural Network Coupled with Rolling Mechanism and Grey Model Algorithms for Streamflow Forecasting Over Multiple Time Horizons",Water Res. Manag, 2018, pg no : 1883–1899

[30] M. A. Chitsazan, M. Sami Fadali, Amanda K. Nelson, A. M Trzynadlowski, "Wind speed forecasting using an echo state network with nonlinear output functions", American Control Conference (ACC), 2017 IEEE, pp. 5306-5311, May. 2017.