Rainfall – Runoff Modeling: Comparison and Combination of Simple Time-Series, Linear Autoregressive and Artificial Neural Network Models

JAGAN MOHAN REDDY, A — Post-graduate Student — ajagan_amohan@yahoo.com
SURESH BABU,Ch — Research Scholar — sureshsuch@gmail.com
MALLIKARJUNA, P — Professor — mallikarjuna_perugu@yahoo.co.in
Department of Civil Engineering
Sri Venkateswara University
Tirupati – 517 502 (A.P.)
INDIA

Abstract: - Runoff simulation and forecasting is essential for planning, designing and operation of water resources projects. In the present study, the rainfall – runoff process is modeled using empirical methods such as simple time - series (STS) and linear autoregressive (ARX) models and compared with Artificial Neural Networks (ANNs). It also explores the improvement in the performance of neural networks by combing them with empirical methods. The study uses the monthly data at Sriramsagar, Mancherial and Polavaram gauging sites of Godavari basin of India. The ANN models, because of their nonlinear modeling capability, outperformed the empirical approaches. The study also reveals that the performance of ANN models in the simulation and forecasting of monthly runoff during monsoon period can be improved considerably by including the residuals derived from STS and ARX models as additional inputs together with rainfall.

Key-Words: - Simple time-series model, Linear autoregressive model, Artificial neural network, Comparison and combination, Performance evaluation

1. Introduction

Accurate estimation of runoff in streams is essential for the development, regulation and efficient management of water resources for different activities such as irrigation, water supply, hydropower development, land drainage and flood control. The process of transformation of rainfall into runoff over a catchment is complex, nonlinear and exhibits both temporal and spatial variability. The models developed to simulate the process can be categorized as empirical, conceptual and physically based distributed models. However, these models fail to represent the inherent nonlinear dynamics in the process of rainfall – runoff transformation. The Artificial Neural Network (ANN) technique which is capable of representing complex nonlinear process has been applied in the recent years, as a successful tool, in the rainfall – runoff modeling.

A number of researchers have investigated the capability of artificial neural networks in the rainfall – runoff modeling process (Halff et al. [8], Zhu et al. [29], Karunanithi et al. [12], Smith and Eli [22], Hsu et al. [9], Minns and Hall [14], Jayawardena and Fernando [11], Shamseldin, [21], Fernando and Jayawardena [6], Dawson and Wilby [4], Tokar and Johnson [24], Campolo et al. [3], Sajikumar and Thandaveswara [19], Tokar and Markus [25], Thirumalaiah and Deo [23], Imrie et al. [10], Elshorbagy [5], Zhang and Govindaraju [28], Birkundavyi et al. [2], Rajurkar et al. [17], Senthil kumar et al.[20], Wu et al. [26], Nilsson et al [16], Garbrecht [7] and Archama Sakar et al. [1]).

In the recent rainfall – runoff modeling studies, different algorithms and transfer functions were proposed to make the ANN training more effective and efficient. The performance of ANN models was improved with addition of input variables and compared with that of conventional models. The models were coupled with black box and conceptual techniques to enhance their modeling capability. The neural based stochastic, geographical information and fuzzy inference systems were also developed for hydrologic prediction. The present study aims to compare the empirical models such as simple linear time - series(STS) and linear autoregressive(ARX) models with ANN models. An attempt is also made to combine these models in order to enhance the forecasting ability of ANN models in the simulation and estimation of monthly runoff during monsoon period at Sriramsagar, Mancherial and Polavaram gauging sites of Godavari basin of India.

2. Study Area and Data Generation

The stream gauging sites selected for the present study are Sriramsagar, Mancherial and Polavaram of Godavari basin of India. A brief description of the gauging sites is presented in Table 1. The locations of influencing rain gauge stations and gauging sites are shown in Fig.1.

Table 1 Brief Description of Gauging Sites

Gauging site	Latitude/ Longitude	Temperature (^{0}C)		Catchment area (km^2)		Mean monsoon rainfall (cm)		Mean monsoon flow (Mm^3)	
		Max.	Min.	Free	Combined	Training period	Testing period	Training period	Testing period
Sriramsagar	18^{0}5'0"N/ 78^{0}20'0"E	45	20	38121	90594	119.8	98.6	11256	10894
Mancherial	18^{0}13'0"N/ 78^{0}30'0"E	43	21	11149	101743	137.0	102.7	11669	8918
Polavaram	17^{0}13'0"N/ 81^{0}46'0"E	44	22	204900	306643	105.9	104.49	70392	83181

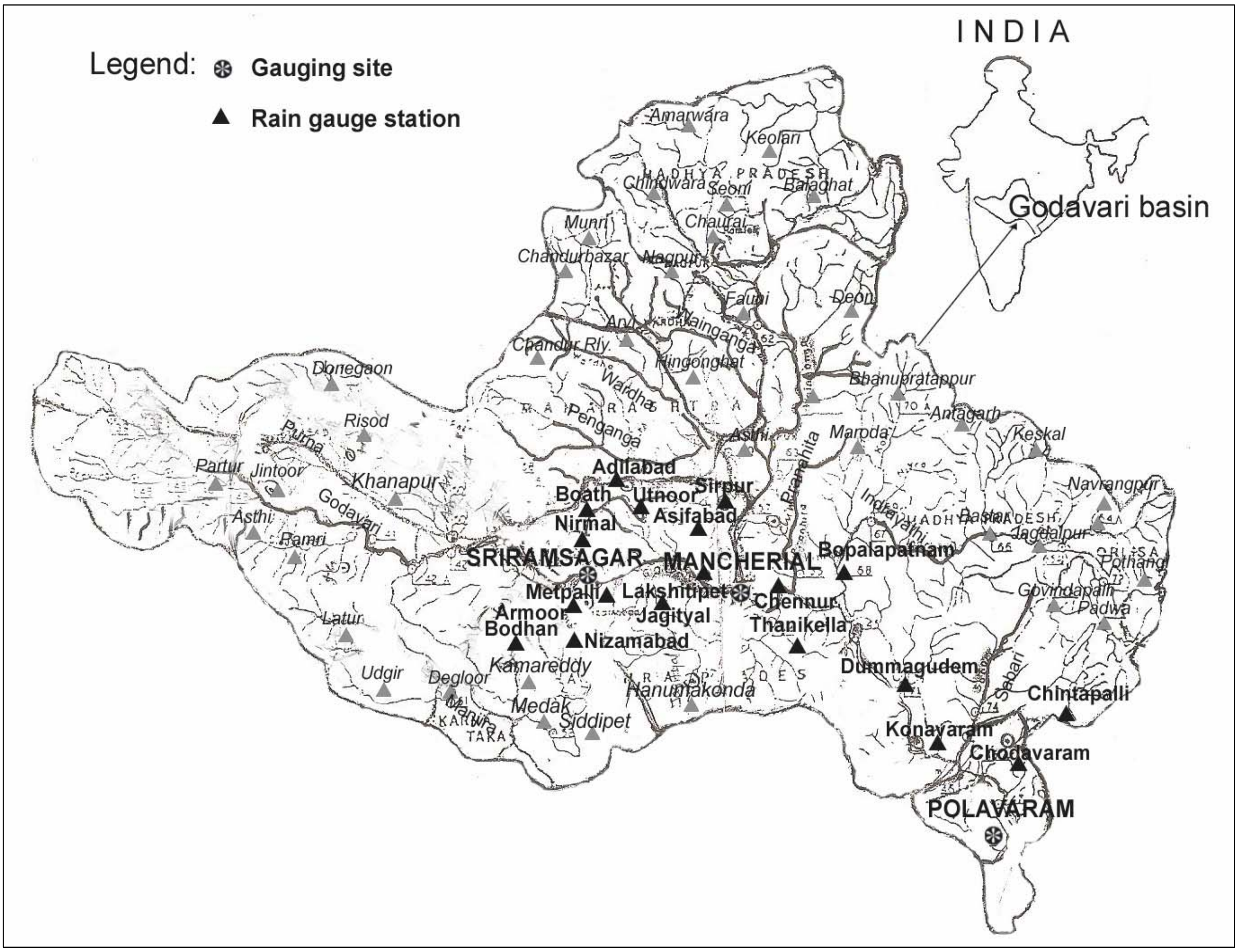

Fig.1: Index Map of Study Area

The monthly rainfall data during monsoon period at rain gauge stations influencing the catchments of gauging sites for the period 1985 – 1999 were collected from the Irrigation department, Andhra Pradesh, India. The corresponding monthly runoff data for the period at gauging sites were obtained from the Central Water Commission (CWC), India. Ten years' data (1985 – 1994) were used for calibration and the rest (1995 – 1999) was used for verification of the models.

3. Methodology

The combination of the following models was adopted in the present study.

3.1 Simple Time-Series Model (STS)

The model is expressed as simple equation based on summation components as

$$R_t = C + \sum_{i=1}^{m} a_i p_{t-i+1} + \sum_{i=1}^{m} b_i R_{t-i} \qquad (1)$$

where a_i and b_i are empirical parameters and P_t and R_t are rainfall and runoff in the time period 't'.

3.2 Linear Autoregressive Model (ARX)

The discrete linear autoregressive model structure is expressed as

$$R_t + \sum_{i=1}^{m} \alpha_i R_{t-i} = \sum_{i=1}^{m} \beta_i P_{t-i} + e_t \qquad (2)$$

where α_i and β_i are modeling parameters, and e_t is the error. The optimum values of the model parameters are estimated using the MATLAB identification toolbox.

3.3 Artificial Neural Network (ANN)

A standard multilayer feed-forward ANN with logistic sigmoid transfer function was adopted for the present study. Error back propagation which was an iterative nonlinear optimization approach based on the gradient descent search method (Rumalhart [18]) was used during calibration. The input data were normalized in the range of (0.1, 0.9) by the equation.

$$x_{norm} = 0.1 + 0.8\,(x_i / x_{max})$$

where x_{norm} is the normalized dimensionless variable, x_i and x_{max} are the observed and maximum values in the data set. The calibration set was used to minimize the error and validation set was used to ensure proper training of the neural network employed such that it does not get over trained. The optimal network was obtained through trial and error process using MATLAB routines.

4. Performance Evaluation Criteria

The evaluation criteria used in the present study are coefficient of determination (R^2), root mean square error (RMSE), efficiency coefficient (EC) and volumetric error (EV).

4.1 Coefficient of Determination (R^2)

It is the square of the correlation coefficient (R) and the correlation coefficient is expressed as

$$R = \frac{\sum_{i=1}^{n} (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\left[\sum_{i=1}^{n} (y_i - \bar{y})^2 \sum_{i=1}^{n} (\hat{y}_i - \bar{\hat{y}})^2\right]^{1/2}} \mathrm{x}100 \qquad (3)$$

where y_i and $\hat{y}_i$ are the observed and estimated values respectively and, $\bar{y}$ and $\bar{\hat{y}}_i$ are the means of observed and estimated values.

4.2 Root Mean Square Error (RMSE)

It yields the residual error in terms of the mean square error expressed as (Yu, [27]).

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n}} \qquad (4)$$

where n = number of observations.

4.3 Efficiency Coefficient (EC)

It is used to assess the performance of different models (Nash and Sutcliffe, [15]). It is a better choice than RMSE statistic when the calibration and verification periods have different lengths (Liang et al.[13]). It is expressed as

$$EC = \left(1 - \frac{F}{F_0}\right)\mathrm{x}100 \qquad (5)$$

where $F_0 = \sum_{i=1}^{n} (y_i - \bar{y})^2$ and $F = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$

A value of EC of 90% generally indicates a very satisfactory model performance while a value in the range 80-90% indicates a fairly good model. Values of EC in the range 60-80% would indicate an unsatisfactory model fit.

4.4 Volumetric Error (EV)

It is an absolute prediction error (Yu, [27]) expressed as

$$EV = \frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)}{\sum_{i=1}^{n} y_i} \mathrm{x}100 \qquad (6)$$

The negative EV values denote underestimation of the output variable.

5. Results and Discussion

The runoff at each of the gauging sites has been estimated using STS, ARX and ANN models independently. The performance of the models was evaluated based on the selected performance evaluation criteria. A comparison of the performance indices of these models is presented in Table 2. It is evident from Table 2 that the ANN models outperformed STS and ARX models in terms of performance indices. However, as the ANNs belong to the class of data driven approaches, it is important to determine the dominant network inputs as this not only reduces the training times but also increases the generalization ability of the network for a given data set. Therefore, the selection of input variables is given utmost importance to define the models. In the present study it was adopted in four ways. In the first case, the rainfall was considered as input variable. In the second case, an additional input of residuals, regressors or runoffs obtained from STS model were considered. In the third case the residuals, regressors or runoffs obtained from ARX model were used while in the fourth case residuals, regressors or runoffs obtained from both STS and ARX models were used as additional inputs.

. The runoff values predicted using ANN models with residuals derived and runoffs estimated from STS and ARX models as additional inputs compared well with the observed values. As the improvement in the performance of ANN models when the residuals derived from STS and ARX models are used as additional inputs is more significant, these results are only presented here.. The performance indices of ANN models with different inputs are presented in Table 3. ANN1, ANN2 and ANN3 models simulate the runoff at Sriramsagar, Mancherial and Polavaram gauging sites respectively with only rainfall as the input variable (Case 1). ANN4, ANN5 and ANN6 simulate the runoff at the gauging sites with residuals derived from STS model as additional input (Case 2). ANN7, ANN8 and ANN9 simulate the runoff with an additional input of residuals derived from ARX model (Case 3). ANN10, ANN11 and ANN12 models use the inputs of rainfall and residuals derived from both STS and ARX models (Case 4).

It may be observed from the results presented in Table 3 that the performance of ANN models with rainfall as the input in terms of R^2 and EC is fairly good. However the performance has improved considerably and satisfactorily with additional inputs of residuals derived from STS and ARX models. The RMSE has also reduced significantly.

Figs. 2 to 9 show the scatter plots of observed and estimated values for the models with different inputs (ANN1 to ANN12) for training and testing periods. It may be observed from Figs. 2 and 6 that there is considerable scatter from the ideal line indicating either overestimation or underestimation of runoff when only rainfall is considered as the input variable. These figures also make clear that the network tends to underestimate the high values during the training period. The possible reason for this tendency in the ANN is due to the small number of high values in the time series that it makes it more difficult for the network to learn such events. The scatter has reduced significantly when the residuals derived from either STS or ARX model are added as an additional input (Figs. 3, 4, 7 and 8). The results have improved further when the residuals derived from both STS and ARX models are used as additional inputs (Figs.5 and 9). The comparisons between the observed and estimated runoffs are shown in Figs. 10 to 12. The figures corresponding to ANN 10, ANN 11 and ANN 12 also depict the closeness of the values and thereby reflect the appropriateness of the modeling technique.

Table 2 Performance Indices of STS,ARX and ANN Models.

Gauging Site	Model	Training data set				Testing data set			
		R^2	RMSE (Mm^3/month)	EC(%)	EV(%)	R^2	RMSE (Mm^3/month)	EC(%)	EV(%)
Sriramsagar	STS	0.11	928.76	11.0	3.43	0.21	1021.23	20.8	-8.41
	ARX	0.23	863.08	22.6	10.91	0.30	988.94	23.5	13.99
	ANN	0.91	301.49	90.5	0.29	0.94	263.98	93.7	-2.42
Mancherial	STS	0.10	1952.88	9.5	-4.87	0.10	2252.28	8.2	9.33
	ARX	0.39	1586.30	39.2	-7.16	0.25	2044.79	23.2	-1.73
	ANN	0.75	1028.00	74.8	-0.29	0.88	774.60	87.1	-7.38
Polavaram	STS	0.26	10718.97	25.7	-2.21	0.18	12805.37	17.5	6.50
	ARX	0.26	10751.04	26.3	3.90	0.48	9188.05	47.6	13.01
	ANN	0.75	6010.00	75.1	4.75	0.69	7238.40	68.5	5.11

Table 3 Performance Indices of ANN Models

Model	Training data set				Testing data set			
	R^2	RMSE (Mm^3/month)	EC(%)	EV(%)	R^2	RMSE (Mm^3/month)	EC(%)	EV(%)
ANN1 (1,2,1)	0.91	301.49	90.52	0.29	0.94	263.98	93.78	-2.42
ANN2 (1,2,1)	0.75	1028.00	74.85	-0.29	0.88	774.6	87.13	-7.38
ANN3 (1,2,1)	0.75	6010.00	75.02	4.75	0.69	7238.40	68.33	5.11
ANN4 (2,4,1)	0.98	120.29	98.00	2.12	0.99	111.13	99.12	0.30
ANN5 (2,4,1)	0.97	353.19	97.00	-2.93	0.99	219.45	96.09	-5.22
ANN6 (2,4,1)	0.96	2302.04	96.00	2.13	0.96	2539.54	95.90	3.42
ANN7 (2,4,1)	0.96	200.52	96.00	0.27	0.98	168.49	97.47	-2.51
ANN8 (2,4,1)	0.94	475.75	94.15	-0.61	0.92	662.18	91.11	9.18
ANN9 (2,4,1)	0.98	1523.55	98.00	-0.84	0.98	1808.9	94.82	-4.56
ANN10 (3,6,1)	0.99	97.39	99.30	-0.82	0.99	98.04	99.26	-1.73
ANN11 (3,6,1)	0.99	321.12	98.50	3.23	0.98	286.41	98.26	1.78
ANN12 (3,6,1)	0.99	1039.51	99.24	0.68	0.99	1373.30	98.57	0.98

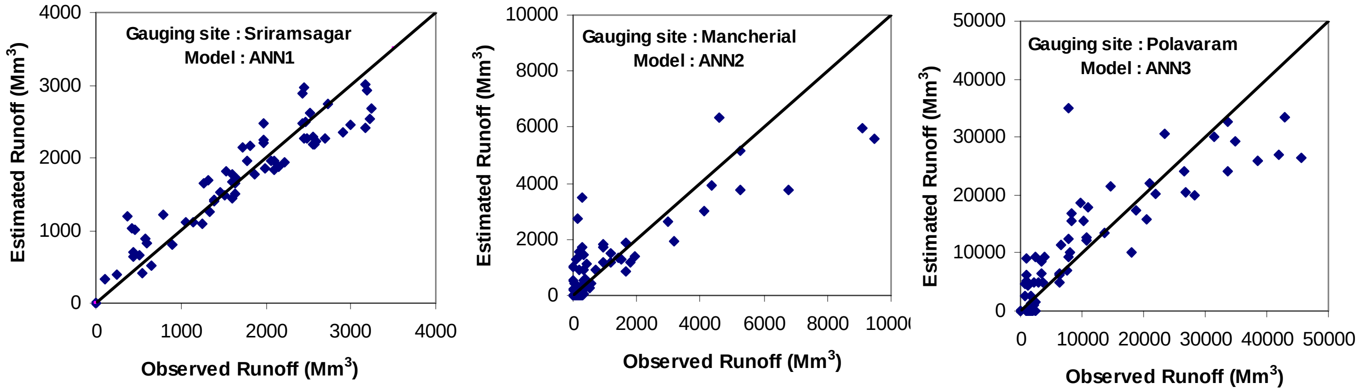

Fig 2 Scatter Plots of Monthly Observed and Estimated Runoff during Training Period (Case 1)

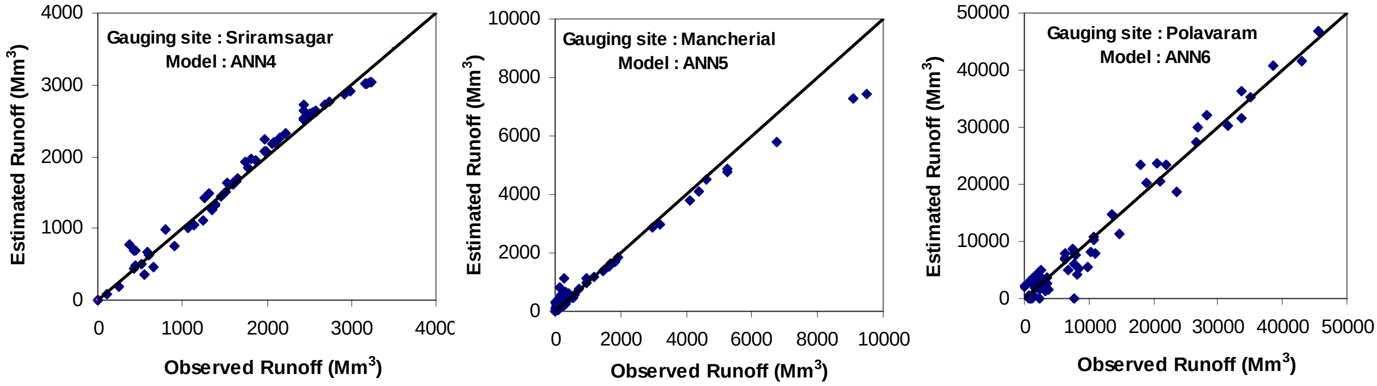

Fig 3 Scatter Plots of Monthly Observed and Estimated Runoff during Training Period (Case 2)

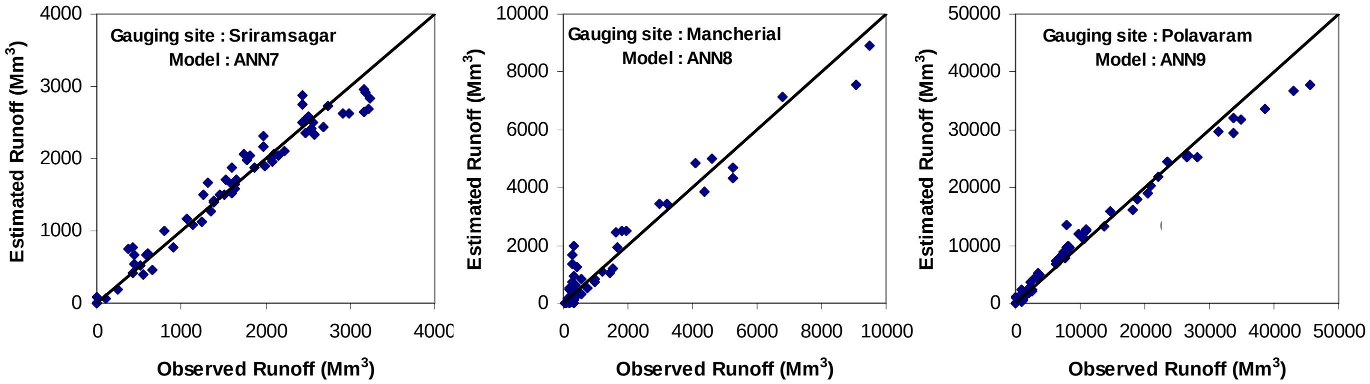

Fig 4 Scatter Plots of Monthly Observed and Estimated Runoff during Training Period (Case 3)

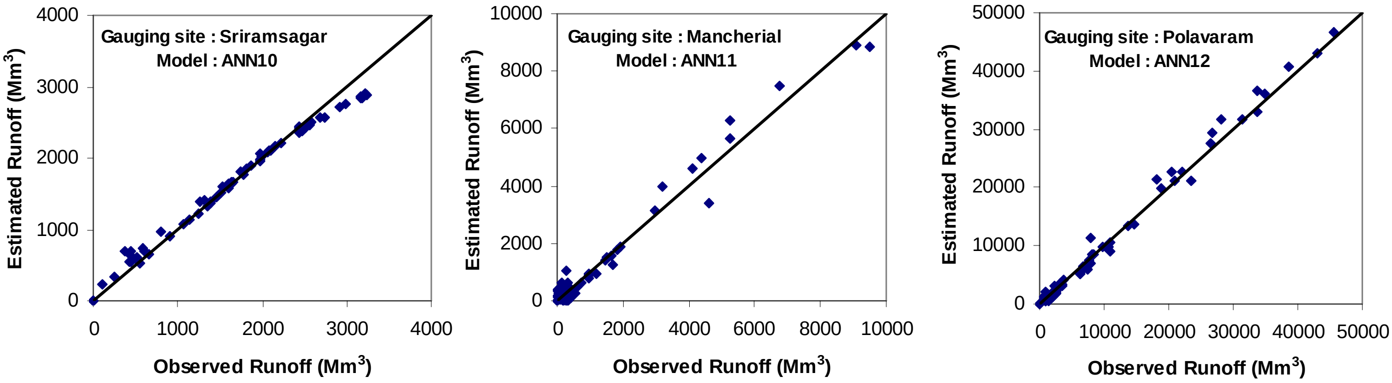

Fig 5 Scatter Plots of Monthly

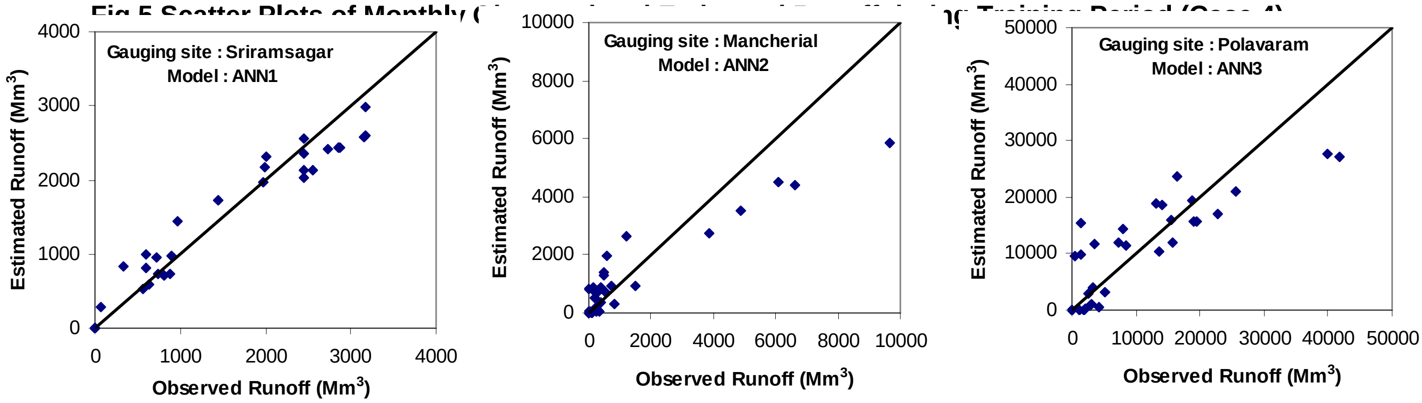

Fig 6 Scatter Plots of Monthly Observed and Estimated Runoff during Testing Period (Case 1)

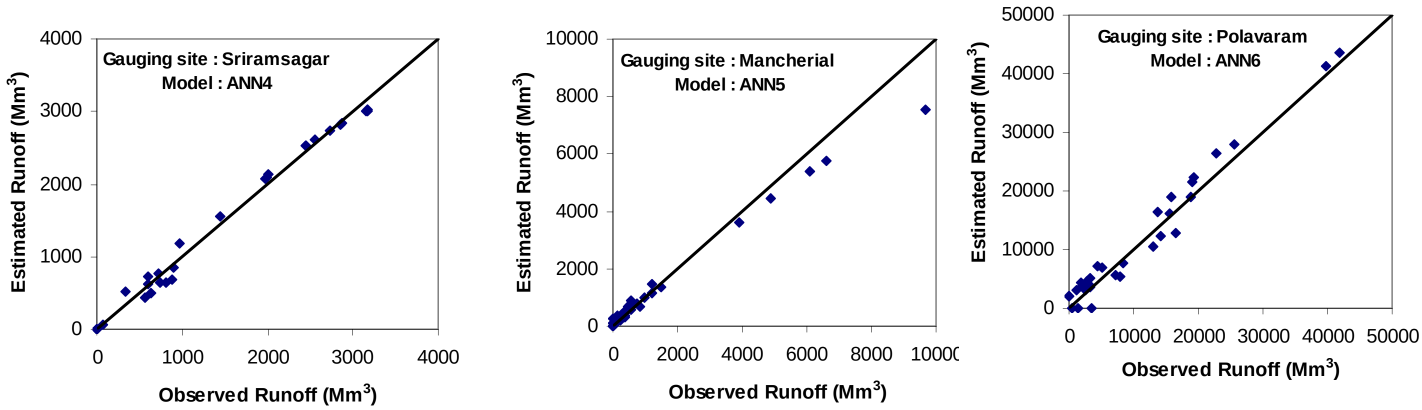

Fig 7 Scatter Plots of Monthly Observed and Estimated Runoff during Testing Period (Case 2)

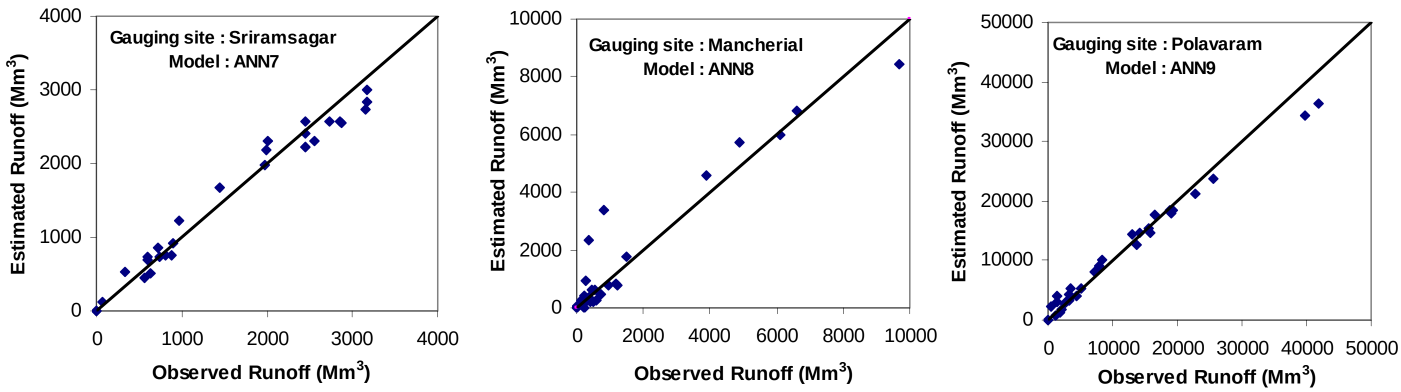

Fig 8 Scatter Plots of Monthly Observed and Estimated Runoff during Testing Period (Case 3)

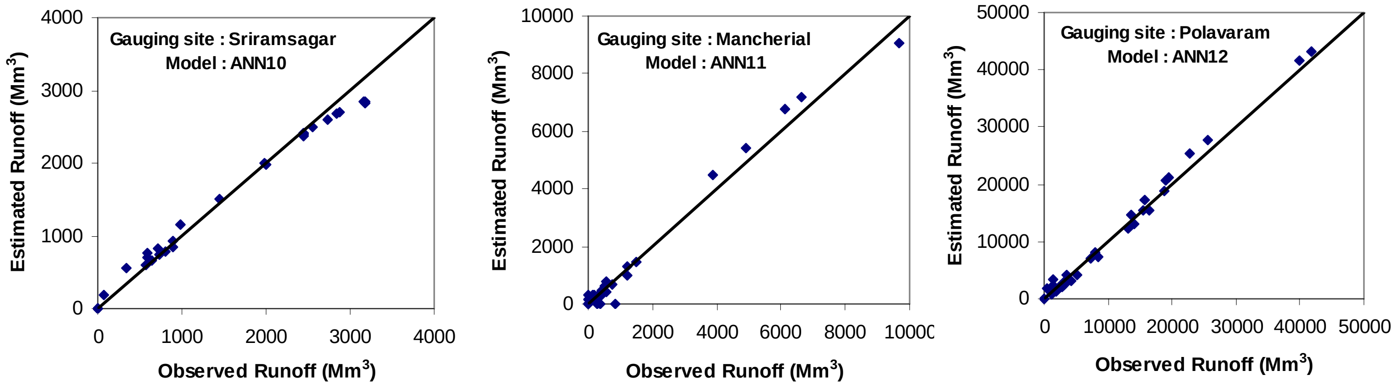

Fig 9 Scatter Plots of Monthly Observed and Estimated Runoff during Testing Period (Case 4)

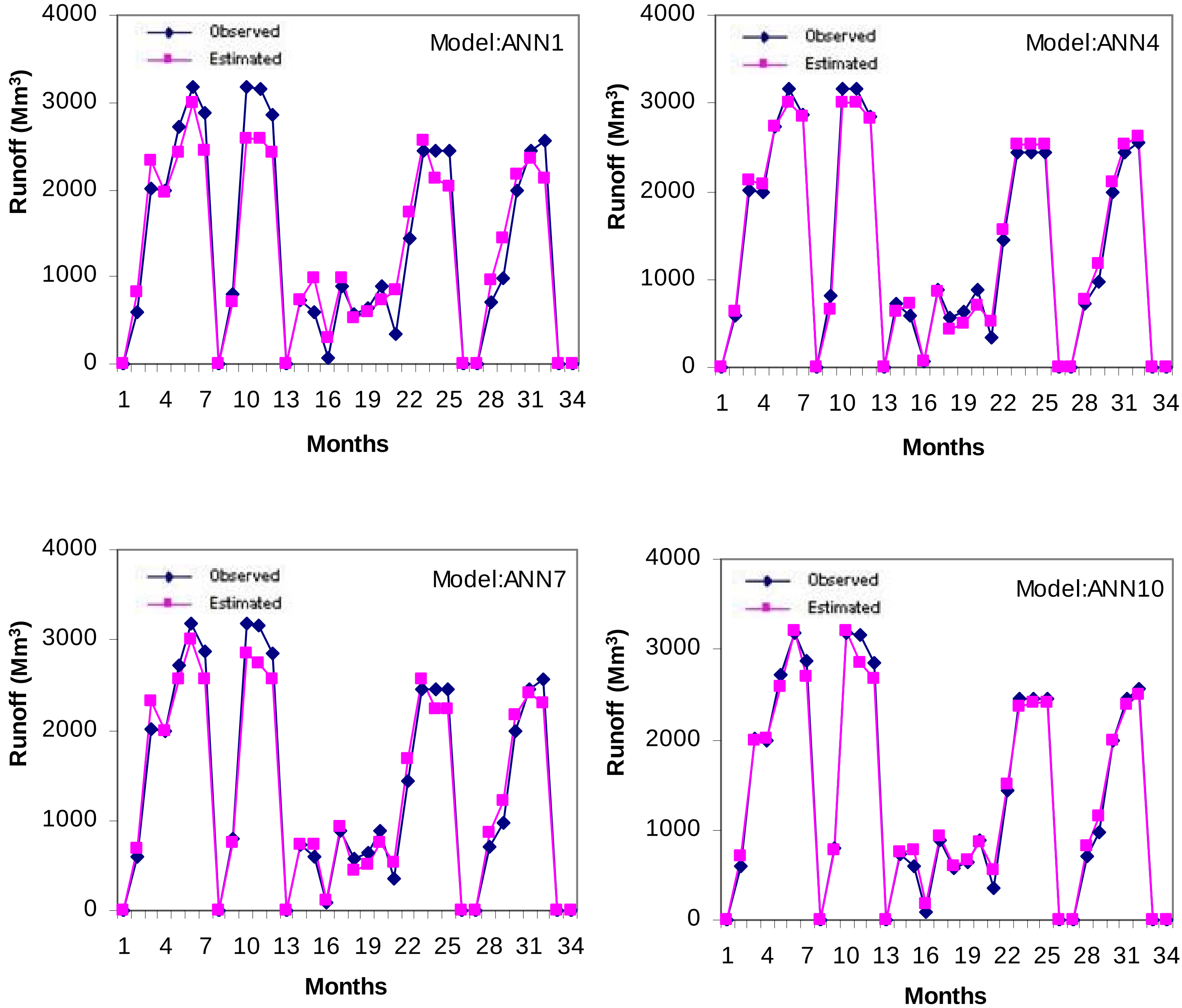

Fig 10 Comparison of Monthly Observed and Estimated Runoff during Testing Period (Gauging Site: Sriramsagar)

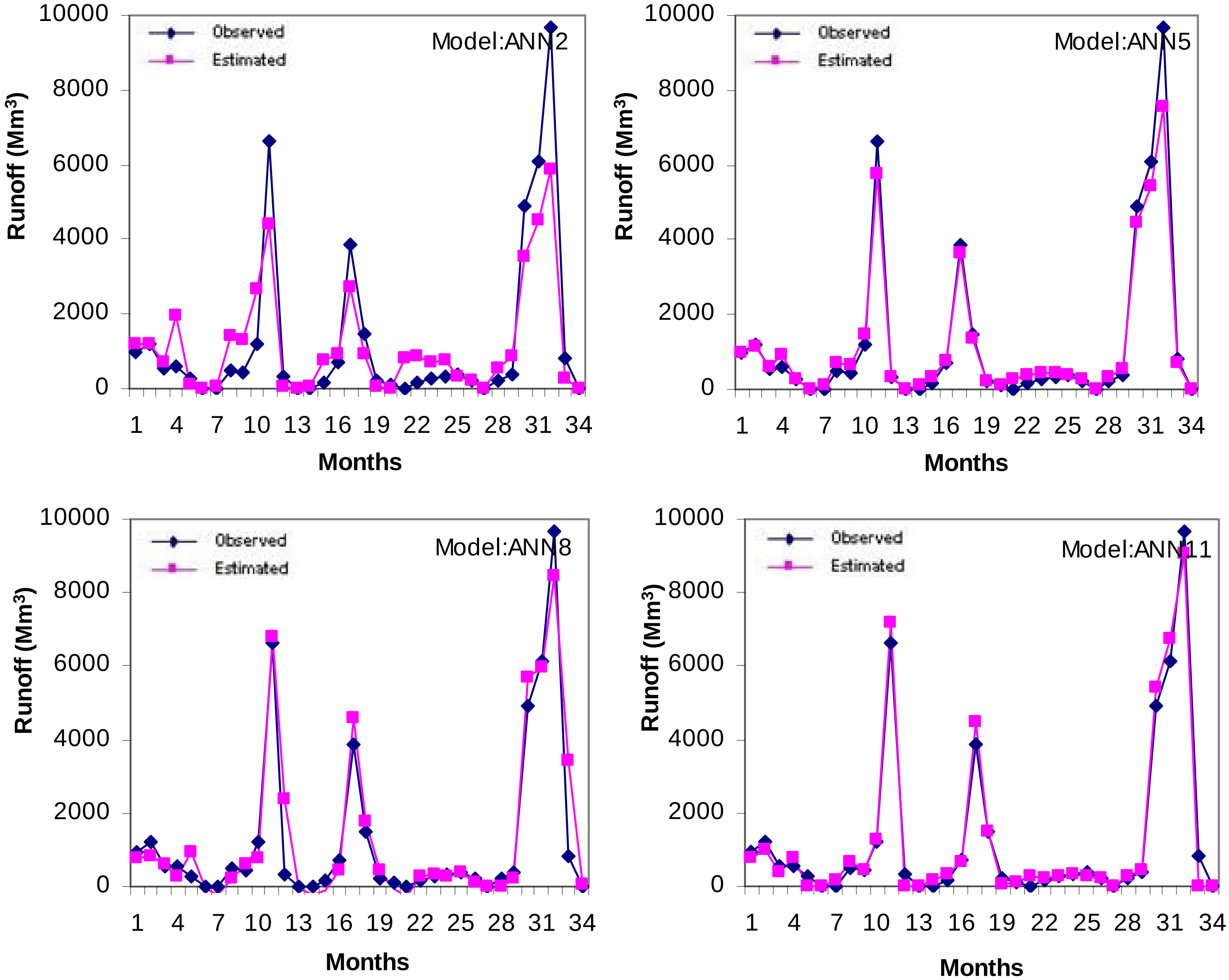

Fig 11 Comparison of Monthly Observed and Estimated Runoff during Testing Period (Gauging Site: Mancherial)

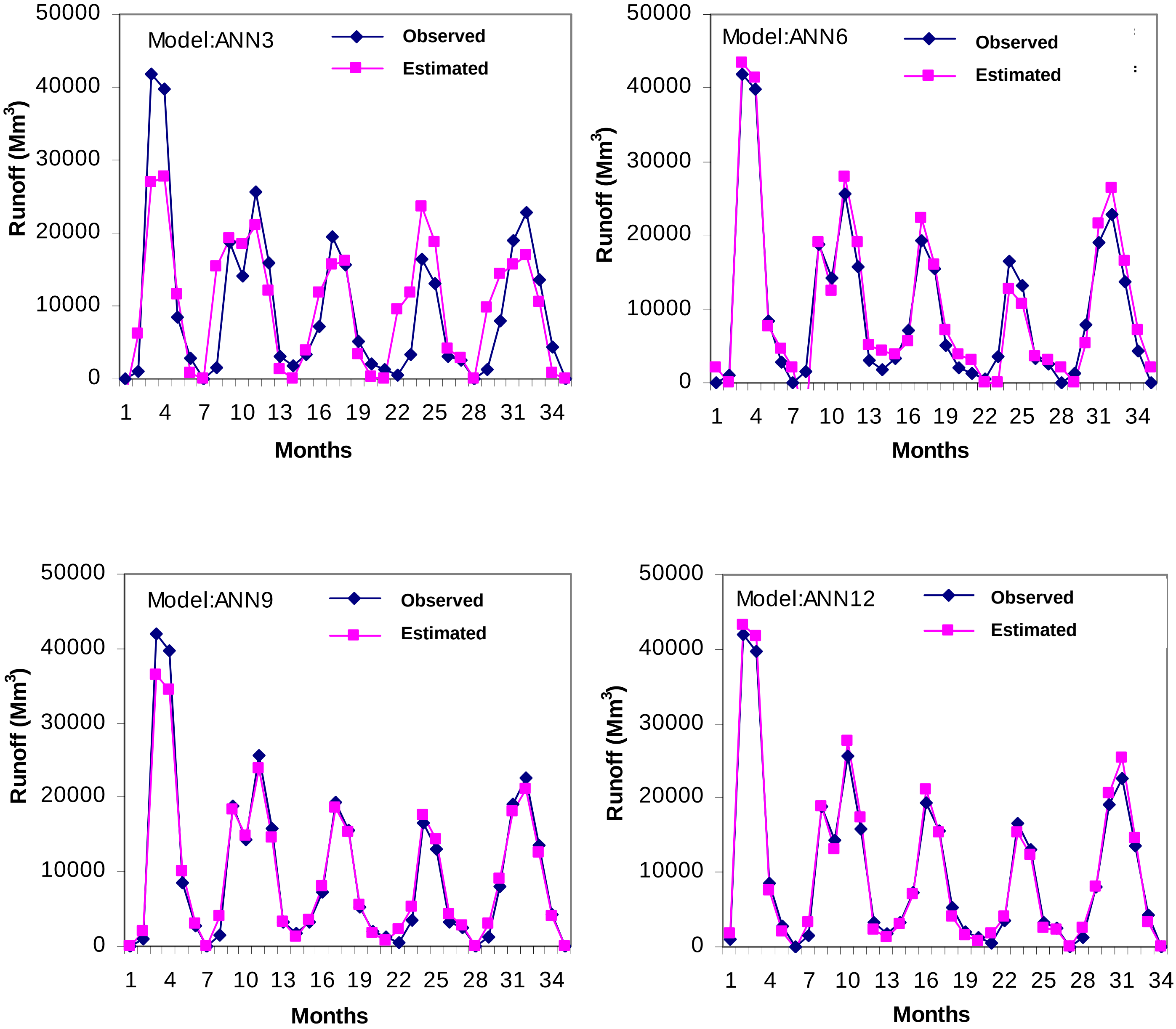

Fig 12 Comparison of Monthly Observed and Estimated Runoff during Testing Period (Gauging Site: Polavaram)

This may be due to the fact that the residuals derived from STS and ARX models which exhibit the nonlinearity present in the rainfall-runoff process reinforce the input to yield satisfactory model output. This modeling technique not only improves the performance but also completes the process with less computational effort (only 100 epochs used). Therefore, the ANN models coupled with STS and ARX models with less computational effort can be adopted for the simulation and forecasting of monthly runoff at the gauging sites of Godavari basin of India.

6. Conclusions

The rainfall – runoff process at Sriramsagar, Mancherial and Polavaram gauging sites of Godavari basin of India is modeled using empirical methods such as simple time - series (STS) and linear autoregressive (ARX) models and compared with Artificial Neural Networks (ANNs). The ANN models, because of their nonlinear modeling capability, outperformed the empirical approaches.

ANN model coupled with simple time - series and linear autoregressive models is also presented. The residuals derived from STS and ARX models are used as additional inputs along with rainfall to improve the performance of ANN model. The results obtained based on the performance evaluation criteria indicate that the ANN model coupled with STS and ARX models can be used satisfactorily in the simulation and forecasting of runoff.

References

[1] Archana Sarkar, Avinash Agarwal, Singh, R.D., Artificial Neural Network Models for Rainfall-Runoff Forecasting in a Hilly Catchment, *Journal of Indian Water Resources Society*, Vol.26, No.3-4, 2006, pp.5-12.

[2] Birikundavyi,S., Labib, R., Rousselle, J., Performance of Neural Networks in Daily Streamflow Forecasting, *Journal of Hydrologic Engineering*, ASCE, Vol.7, No.5, 2002, pp.392-398.

[3] Campolo, M., Andreussi, P., Soldati, A., River Food Forecasting with a Neural Network Model, *Water Resources Research*, Vol.35, No.4, 1999, pp.1191-1197.

[4] Dawson C.W., Wilby, R., An Artificial Neural Network Approach to Rainfall – Runoff Modeling, *Journal of. Hydrological Sciences*, Vol.43, No.1, 1998, pp.47 – 66.

[5] Elshorbagy, A., Simonovic, S.P., Panu, U.S., Performance Evaluation of Artificial Neural Networks for Runoff Prediction, *Journal of Hydrologic Engineering*, ASCE, Vol.5, No.4, 2000, pp.424 – 427.

[6] Fernando, D.A.K., Jayawardena, A.W., Runoff Forecasting using RBF Networks with OLS Algorithm, *Journal Hydrologic Engineering*, ASCE, Vol.3, No.3, 1998, pp.203 – 209.

[7] Garbrecht, J.D., Comparison of Three Alternative ANN Designs for Monthly Rainfall-Runoff Simulation, *Journal of Hydrologic Engineering*, ASCE, Vol.11, No.5, 2006, pp.502-505.

[8] Halff, A.H., Halff, H.M., Azmoodeh, M., Predicting Runoff from Rainfall using Neural Networks, *Proc.*, *Engineering Hydrology*, ASCE, New York, 1993, pp.760 – 765.

[9] Hsu, K., Gupta, H.V., Sorooshian. S., Artificial Neural Network Modeling of the Rainfall - Runoff process, *Water Resources Research*, Vol.31, No.10, 1995, pp.2517 – 2530.

[10] Imrie,C.E., Durcan,S., Korre,A., River flow Prediction using Artificial Neural Networks: Generalization Beyond the Calibration Range, *Journal of Hydrology*, Vol. 233, 2000, pp.138-153.

[11] Jayawardena, A.W., Fernando, D.A.K., Comparison of Multi – Layer Perceptron and Radial Basis Function Network as Tools for Flood Forecasting. *Proc.*, *North American Water and Environment Conference*, ASCE, New York, 1996, pp.457 – 458.

[12] Karunanithi, N., Grenney, W.J., Whitley, D., Bovee, K., Neural Networks for River Flow Prediction, *Journal of Computing in Civil Engineering*, ASCE, Vol.8, No.2, 1994, pp.201 – 220.

[13] Liang, G.C.,O'Connor, K.M., Kachroo, R.K., A Multiple Input Single Output Variable Gainfactor Model, *Journal of Hydrology*., Vol.155, 1994, pp.185 – 198.

[14] Minns, A.W., Hall, M.J., Artificial Neural Networks as Rainfall – Runoff Models, *Journal of Hydrological Sciences*, Vol.41, No.3, 1996, pp.399 – 418.

[15] Nash, J.E, Sutcliffe, J.V., River Flow Forecasting through Conceptual Models I: A Discussion of Principles, *Journal of Hydrology*, Vol.10, No.3, 1970, pp.238 – 250.

[16] Nilsson,P., Cintia,B.U., Berndtsson, R., Monthly Runoff Simulation: Comparing and Combining Conceptual and Neural Network Models, *Journal of Hydrology*, Vol. 321, 2006, pp.344-363.

[17] Rajurkar, M.P., Kothyari, U.C., Chaube, U.C., Modeling of the Daily Rainfall-Runoff Relationship with Artificial Neural Network, *Journal of Hydrology*, Vol. 285, 2004, pp.96-113.

[18] Rumelhart,D.E., Hinton,G.E., Williams, R.J., Learning Internal Representations by Error Propagation, *Parallel Distributed Processing*, Vol.I, MIT Press, Cambridge, MA,1986,pp.318-362.

[19] Sajikumar, N., Thandaveswara, B.S., A Nonlinear Rainfall - Runoff Model Using Artificial Neural Network, *Journal of Hydrology*, Vol.216, 1999, pp.32 – 55.

[20] Senthil Kumar, A.R., Sudheer, K.P., Jain,S.K., Agarwal, P.K., Rainfall-Runoff Modeling using Artificial Neural Networks: Comparison of Network Types, *Journal of Hydrological Processes*, Vol. 19, 2005, pp.1277-1291.

[21] Shamseldin, A.Y., Application of a Neural Network Technique to Rainfall-Runoff Modeling, *Journal of Hydrology*, Vol.199, 1997, pp.272-294.

[22] Smith, J., Eli., R.N., Neural Network Models of Rainfall-Runoff Process, *Journal of Water Resources Planning and Management*, ASCE, Vol.121, No.6, 1995, pp.499-508.

[23] Thirumalaiah, K., Deo, M.C., Hydrological Forecasting using Neural Networks, *Journal of Hydrologic Engineering*, ASCE, Vol.5, No.2, 2000, pp.180-189.

[24] Tokar, A.S., and Johnson, P.A., Rainfall-Runoff Modeling using Artificial Neural Networks, *Journal of Hydrologic Engineering*, ASCE, Vol.4, No.3, 1999, pp.232-239.

[25] Tokar, A.S., Markus, M., Precipitation – Runoff Modeling using Artificial Neural Networks and Conceptual Models, *Journal of Hydrologic Engineering*, ASCE, Vol.5, No.2, 2000, pp.156-161.

[26] Wu, J.S., Han, J., Annambhotla, S., Bryant, S., Artificial Neural Networks for Forecasting Watershed Runoff and Streamflows, *Journal of Hydrologic Engineering*, Vol.10, No.3, 2005, pp.216 – 222.

[27] Yu, P.S., Liui. C.L., Lee., T.Y., Application of Transfer Function Model to a Storage Runoff Process, *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*, Vol.3, 1994, pp.87 – 97.

[28] Zhang, B., Govindaraju, R.S., Prediction of Watershed Runoff using Bayesian Concepts and Modular Neural Networks, *Water Resources Research*, Vol. 36, No.3, 2000, pp.753 – 762.

[29] Zhu, M., Fujita, M., Hashimoto, N., Application of Neural Networks to Runoff Prediction, *Stochastic and Statistical Methods in Hydrology and Environmental Engineering*, Vol.3, K.W.Hipel et al., eds., Kluwer, Dordrecht, The Netherlands, 1994, pp.205-216.