International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 04, April-2017

# Identification of the Input Vector for R-R Modelling of River Jhelum Catchment

Lateef Ahmad Dar
National Institute of Technology,
Srinagar, J&K.

***Abstract:*** **Hydrologic engineering design and management purposes require information about runoff from a hydrologic catchment. In order to predict this information, the transformation of rainfall on a catchment to runoff from it must be modelled. One approach to this modelling issue is to use empirical Rainfall-Runoff (R-R) models. Artificial neural networks (ANNs) are among the most sophisticated empirical models available and have proven to be especially good in modelling complex systems. Their ability to extract relations between inputs and outputs of a process, without the physics being explicitly provided to them, theoretically suits the problem of relating rainfall to runoff well, since it is a highly nonlinear and complex problem. The goal of this investigation was to develop rainfall-runoff models for the river Jhelum catchment that are capable of accurately modelling the relationships between rainfall and runoff in a catchment. It is for this reason that ANN and MLR techniques were tested as R-R models on a data set from the upper Jhelum catchment in Jammu and Kashmir, India. For modeling the rainfall-runoff process in river Jhelum ,the input i.e. the precipitation was determined by taking the data from three rainguage stations viz. Srinagar, Pahalgam and Qazigund for years 2001-2013. The runoff data was taken for padshahibagh guaging station that lies in Srinagar, summer capital of Jammu and Kashmir for the years 2001-2013. From the predicative analysis , it was found that the flow at the Padshahibagh on any given day is dependent on previous three day flows .The number of previous day rainfall inputs influencing the discharge was determined by the trial and error method, the number of previous day rainfall inputs were increased from one to five. The comparison was based on various statistical parameters like root mean square error (RMSE) and $R^2$.**

***Keywords: Rainfall; Runoff; Multiple linear regression (MLR) and Artificial neural network (ANN) technique.***

## I.INTRODUCTION

Rainfall-Runoff relationship remains one of the most complex processes in hydrological analysis. Rainfall runoff modelling is a very important tool for determining runoff generated for a particular amount of rainfall over the catchment. The response of a catchment to any amount of rainfall occurring over it is influenced by the precipitation characteristics as well as the catchment characteristics. There are a lot of things that influence the response of the catchment like rainfall depth, rainfall intensity, temperature , wind-velocity, slope of the catchment, drainage density, type of soil, vegetative cover, topography etc. to determine the exact influence of these parameters on the overall response of the catchment is a very complex and cumbersome phenomena. We have to find and incorporate only those parameters which have a considerable influence on the response of the catchment and hence save the time spent on analysing the non-important variables. To find the parameters that influence the output considerably, we have to use the predicative analysis technique. The present study involves determining the inputs that are to be used in modelling the rainfall-runoff process of river Jhelum catchment using black-box techniques Viz. Multiple linear regression (MLR) and Artificial neural network (ANN) technique.

## II.STUDY AREA

The present study was done on the upper Jhelum catchment. The study area spatially lies between 33° 21′ 54″ N to 34° 27′ 52″ N latitude and 74° 24′ 08″ E to 75° 35′ 36″ E longitude with a total area of 8600.78 sq.kms (Fig.1). It covers almost all the physiographic divisions of the Kashmir Valley and is drained by the most important tributaries of river Jhelum. Srinagar city which is the largest urban centre in the valley is settled on both the sides of Jhelum River and is experiencing a fast spatial growth. Physical features of contrasting nature can be observed in the study area that ranges from fertile valley floor to snow-clad mountains and from glacial barren lands to lush green forests. Based on the consistency of data, the precipitation data was taken from three rain-gauge stations in the catchment viz. Srinagar, Qazigund and Pahalgam. The discharge was taken at Padshahibagh gauging station. The study area can broadly be divided into the following physiographic divisions viz Mountainous region of Pir Panjal and Greater Himalayas,The lacustrine deposits of Karewa , Jhelum Valley Floor. The geological history of the study area ranges from Cambrian to Recent. The central alluvial part of the Upper Jhelum catchment is a Recent formation surrounded by Karewas on the south and south-west and Jurassic formations on the north north-east and north-west. These three are the major formations found in the study area. South eastern part of the study area is composed of Triassic and Cambro- Silurian formations. Few linear stretches in the north of the study area are of Triassic and Jurassic formations interspersed with unclassified granites and gneisses. The Jhelum and its associated streams that drain the bordering mountain slopes together constitute the drainage network of the study area. They include the fairly developed systems of the Sind, Rembiara, Vishaw and Lidder rivers as well as tiny rivulets such as the Sandran, Bringi and Arapat Kol . Adjusted to the varying nature of geomorphic and geological setting, the fluvial systems in the study area have peculiar characteristics of their own. Drainage system of the Upper Jhelum catchment has an evolutionary history marked by stupendous changes in level, rejuvenating at one time, and at others becoming sluggish, or even choking their channels with their own debris with consequent diversions and the ever-threatening process of mutual piracy.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 04, April-2017

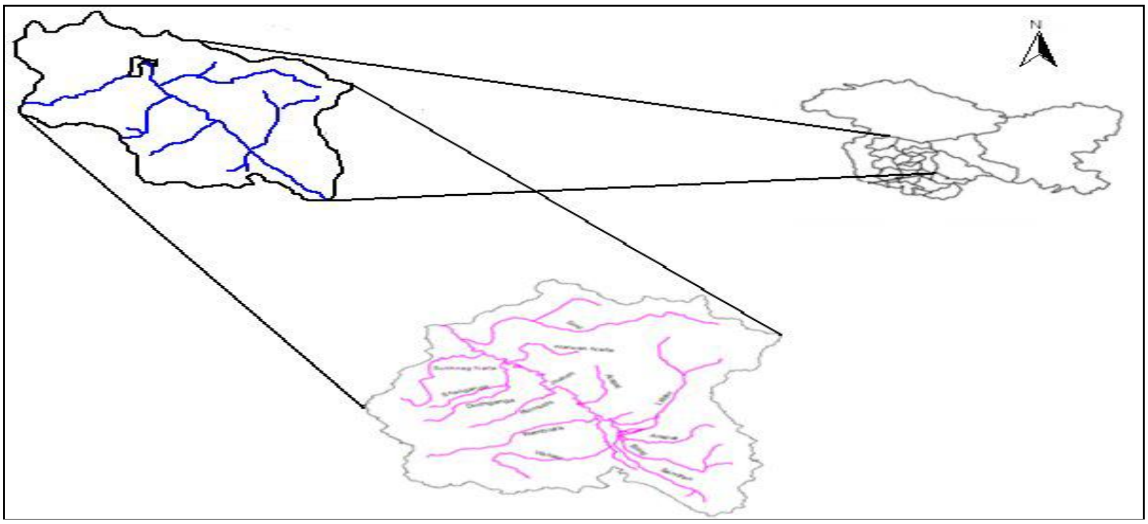

Figure.1:Study area (Source: Generated from SOI toposheets, 1961)

## III. IDENTIFICATION OF THE INPUT VECTOR

Identification of the number of flow series was carried by the predicative analysis. The selection of the predictors was carried out on the basis of p-test.

| | A | B | C | D | E | F | G | H | I | J | K |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Q(T) | Q(T-1) | Q(T-2) | Q(T-3) | Q(T-4) | Q(T-5) | Q(T-6) | Q(T-7) | Q(T-8) | Q(T-9) | Q(T-10) |
| 2 | 126.58 | | | | | | | | | | |
| 3 | 133.01 | 126.5783 | | | | | | | | | |
| 4 | 136.55 | 133.0144 | 126.5783 | | | | | | | | |
| 5 | 139.36 | 136.5509 | 133.0144 | 126.5783 | | | | | | | |
| 6 | 139.36 | 139.3609 | 136.5509 | 133.0144 | 126.5783 | | | | | | |
| 7 | 140.76 | 139.3609 | 139.3609 | 136.5509 | 133.0144 | 126.5783 | | | | | |
| 8 | 144.24 | 140.7597 | 139.3609 | 139.3609 | 136.5509 | 133.0144 | 126.5783 | | | | |
| 9 | 150.44 | 144.2392 | 140.7597 | 139.3609 | 139.3609 | 136.5509 | 133.0144 | 126.5783 | | | |
| 10 | 174.56 | 150.4414 | 144.2392 | 140.7597 | 139.3609 | 139.3609 | 136.5509 | 133.0144 | 126.5783 | | |
| 11 | 237.71 | 174.5554 | 150.4414 | 144.2392 | 140.7597 | 139.3609 | 139.3609 | 136.5509 | 133.0144 | 126.5783 | |
| 12 | 188.82 | 237.7087 | 174.5554 | 150.4414 | 144.2392 | 140.7597 | 139.3609 | 139.3609 | 136.5509 | 133.0144 | 126.5783 |
| 13 | 167.96 | 188.8172 | 237.7087 | 174.5554 | 150.4414 | 144.2392 | 140.7597 | 139.3609 | 139.3609 | 136.5509 | 133.0144 |
| 14 | 143.55 | 167.9601 | 188.8172 | 237.7087 | 174.5554 | 150.4414 | 144.2392 | 140.7597 | 139.3609 | 139.3609 | 136.5509 |
| 15 | 136.55 | 143.5453 | 167.9601 | 188.8172 | 237.7087 | 174.5554 | 150.4414 | 144.2392 | 140.7597 | 139.3609 | 139.3609 |
| 16 | 131.59 | 136.5509 | 143.5453 | 167.9601 | 188.8172 | 237.7087 | 174.5554 | 150.4414 | 144.2392 | 140.7597 | 139.3609 |
| 17 | 142.15 | 131.5922 | 136.5509 | 143.5453 | 167.9601 | 188.8172 | 237.7087 | 174.5554 | 150.4414 | 144.2392 | 140.7597 |
| 18 | 145.62 | 142.1545 | 131.5922 | 136.5509 | 143.5453 | 167.9601 | 188.8172 | 237.7087 | 174.5554 | 150.4414 | 144.2392 |
| 19 | 137.25 | 145.6241 | 142.1545 | 131.5922 | 136.5509 | 143.5453 | 167.9601 | 188.8172 | 237.7087 | 174.5554 | 150.4414 |
| 20 | 136.55 | 137.255 | 145.6241 | 142.1545 | 131.5922 | 136.5509 | 143.5453 | 167.9601 | 188.8172 | 237.7087 | 174.5554 |
| 21 | 136.55 | 136.5509 | 137.255 | 145.6241 | 142.1545 | 131.5922 | 136.5509 | 143.5453 | 167.9601 | 188.8172 | 237.7087 |
| 22 | 181.73 | 136.5509 | 136.5509 | 137.255 | 145.6241 | 142.1545 | 131.5922 | 136.5509 | 143.5453 | 167.9601 | 188.8172 |
| 23 | 310.75 | 181.7271 | 136.5509 | 136.5509 | 137.255 | 145.6241 | 142.1545 | 131.5922 | 136.5509 | 143.5453 | 167.9601 |
| 24 | 268.31 | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 7 | Standard E | 19.64804 | | | | | | | |
| 8 | Observatic | 36 | | | | | | | |
| 9 | | | | | | | | | |
| 10 | ANOVA | | | | | | | | |
| 11 | | *df* | *SS* | *MS* | *F* | *ignificance F* | | | |
| 12 | Regressior | 14 | 16810.36 | 1200.74 | 3.110358 | 0.009344 | | | |
| 13 | Residual | 21 | 8106.958 | 386.0456 | | | | | |
| 14 | Total | 35 | 24917.32 | | | | | | |
| 15 | | | | | | | | | |
| 16 | | *Coefficients* | *andard Err* | *t Stat* | *P-value* | *Lower 95%* | *Upper 95%* | *ower 95.0%* | *pper 95.0%* |
| 17 | Intercept | 31.37667 | 31.26511 | 1.003568 | 0.327011 | -33.6427 | 96.39602 | -33.6427 | 96.39602 |
| 18 | X Variable | 0.738124 | 0.205672 | 3.588835 | 0.001729 | 0.310405 | 1.165843 | 0.310405 | 1.165843 |
| 19 | X Variable | -0.17307 | 0.252278 | -0.68601 | 0.005002 | -0.69771 | 0.351575 | -0.69771 | 0.351575 |
| 20 | X Variable | -0.03293 | 0.253339 | -0.12998 | 0.008978 | -0.55978 | 0.493917 | -0.55978 | 0.493917 |
| 21 | X Variable | -0.21651 | 0.250913 | -0.86289 | 0.079794 | -0.73831 | 0.305291 | -0.73831 | 0.305291 |
| 22 | X Variable | 0.505246 | 0.404672 | 1.24853 | 0.225577 | -0.33632 | 1.346808 | -0.33632 | 1.346808 |
| 23 | X Variable | -0.09586 | 0.42605 | -0.22501 | 0.824149 | -0.98188 | 0.790155 | -0.98188 | 0.790155 |
| 24 | X Variable | -0.20219 | 0.424291 | -0.47654 | 0.638608 | -1.08456 | 0.680169 | -1.08456 | 0.680169 |
| 25 | X Variable | -0.23752 | 0.427611 | -0.55545 | 0.584458 | -1.12678 | 0.65175 | -1.12678 | 0.65175 |
| 26 | X Variable | 0.170074 | 0.430264 | 0.395278 | 0.696624 | -0.72471 | 1.064858 | -0.72471 | 1.064858 |
| 27 | X Variable | 0.291467 | 0.429911 | 0.677971 | 0.505192 | -0.60258 | 1.185517 | -0.60258 | 1.185517 |
| 28 | X Variable | 0.029067 | 0.433959 | 0.06698 | 0.947231 | -0.8734 | 0.931533 | -0.8734 | 0.931533 |
| 29 | X Variable | -0.1089 | 0.35194 | -0.30944 | 0.760037 | -0.8408 | 0.622995 | -0.8408 | 0.622995 |
| 30 | X Variable | -0.2143 | 0.293119 | -0.73109 | 0.472803 | -0.82387 | 0.395278 | -0.82387 | 0.395278 |
| 31 | X Variable | 0.403096 | 0.239358 | 1.684073 | 0.106975 | -0.09468 | 0.900867 | -0.09468 | 0.900867 |

Figure 2: Predictative analysis for the number of input flow series.

The predicative analysis suggests incorporating flow values with three days lag in the input vector to the network. Table 1 shows the results obtained from the predicative analysis.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 04, April-2017

Table 1: Results of P-Test.

| Parameter | P-value |
|---|---|
| Intercept | 0.327011116 |
| Q(t-1) | 0.001728645 |
| Q(t-2) | 0.005002098 |
| Q(t-3) | 0.008978154 |
| Q(t-4) | 0.079794358 |
| Q(t-5) | 0.225577479 |
| Q(t-6) | 0.824148923 |
| Q(t-7) | 0.638607708 |
| Q(t-8) | 0.584458284 |
| Q(t-9) | 0.696624304 |
| Q(t-10) | 0.505192364 |
| Q(t-11) | 0.94723135 |
| Q(t-12) | 0.760036923 |
| Q(t-13) | 0.472803194 |
| Q(t-14) | 0.106974693 |

The p-value of Q(t-1), Q(t-2), Q(t-3) is less than 0.05 while all other flow inputs from Q(t-4) onwards exceed P-value of 0.05 and hence fail this test. So previous three day flows will influence the present day flow at Padshahi-Bagh. So previous three day flows values are to be put in the input vector.

Hence the predicative analysis suggests incorporating flow values with three days lag in the input vector to the network.

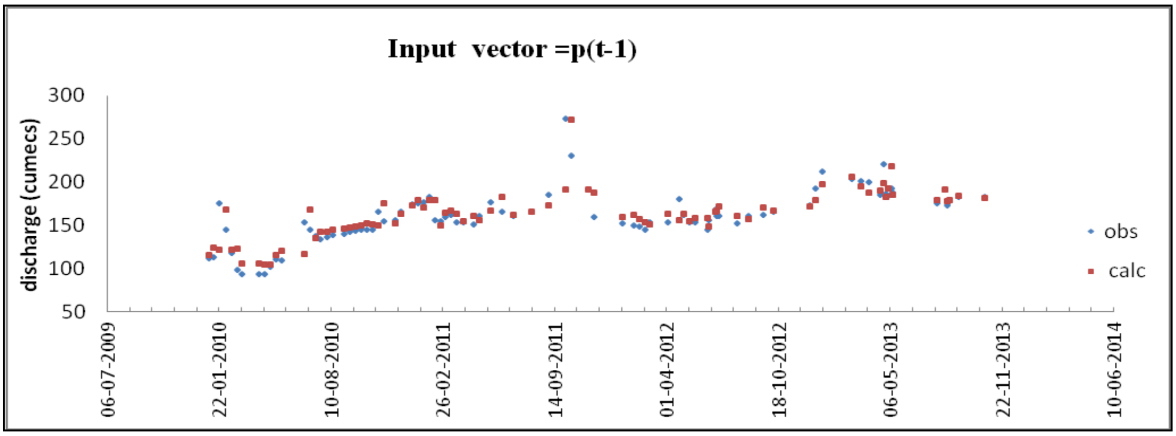


Figure 3: Observed vs. predicted discharge : input P(t-1).

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 6 Issue 04, April-2017

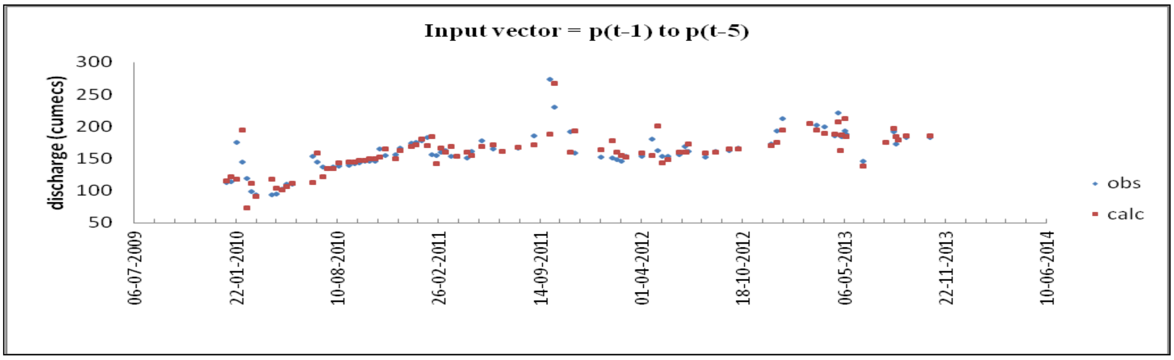


Figure.4: Observed vs. predicted discharge input P(t-1)…P(t-5).

Table 2: Statistics of MLR.

| INPUT | $R^2$ | MSE | RMSE |
|---|---|---|---|
| p(T-1) | 0.764 | 2.435 | 1.560 |
| p(T-1)… p(T-2) | 0.818 | 1.474 | 1.214 |
| p(T-1)…p(T-3) | 0.823 | 0.934 | 0.967 |
| p(T-1)… p(T-4) | 0.825 | 0.465 | 0.681 |
| p(T-1) …p(T-5) | 0.827 | 0.234 | 0.483 |

The RMSE and $R^2$ values improve considerably on increasing the number of rainfall inputs from one to five.

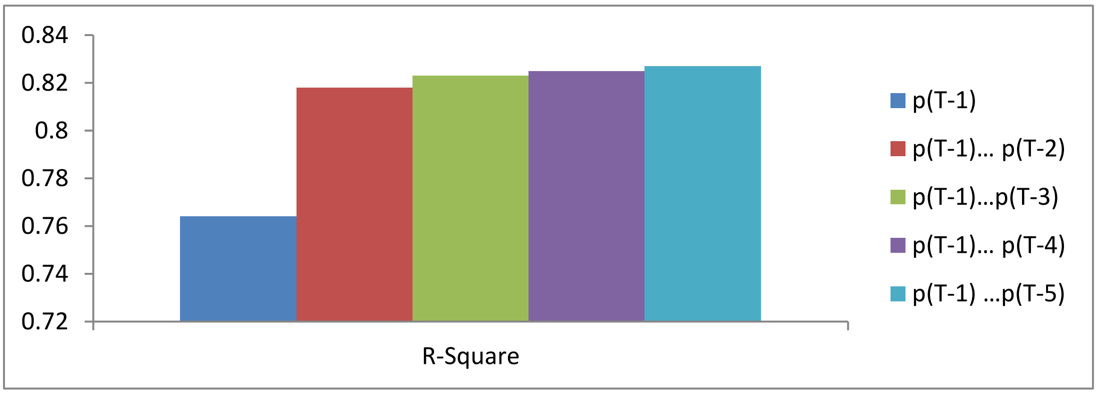


Figure 5: Variation of $R^2$ with different input vectors in MLR

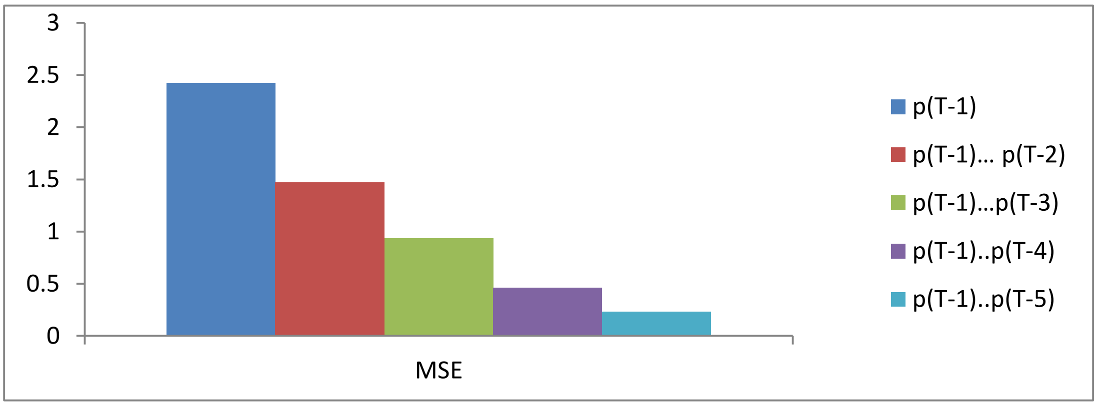


Figure 6: Variation of RMSE with different input vectors in MLR.

Table 3: Goodness of fit for the effect of no. of previous day input rainfall parameters.

| NETWORK | INPUTS | RMSE | R-SQUARE |
|---|---|---|---|
| Back propagation network 6 neurons. | P(t-1) | 0.521 | 0.776 |
| | P(t-1)…P(t-5) | 0.245 | 0.856 |
| Back propagation network 10 neurons. | P(t-1) | 0.342 | 0.814 |
| | P(t-1)…P(t-5) | 0.132 | 0.891 |

The RMSE error improves when the number of input rainfall patterns in the input vector is increased from one to five. The first case considered 4 variables in the input vector while the second considered eight. The above results conclude that the number of previous rainfall data has a significant effect on the model performance. Hence the input vector with 1, 2, 3, 4, 5 day lag can produce the river flow patterns in a satisfactory manner.

Table 4: Statistical indices of RBF model for various inputs.

| NETWORK | INPUTS | RMSE | R-SQUARE |
|---|---|---|---|
| RBF | P(t-1) | 0.097 | 0.912 |
| | P(t-1)…P(t-5) | 0.046 | 0.937 |

The statistical analysis shows that the performance of this network increases as the number rainfall inputs increase from one to five i.e. from p(t-1) to p(t-1)....p(t-5).

## IV.CONCLUSIONS

From the predicative analysis it can be concluded that the discharge at Padshahi-Bagh is dependent upon previous three days flow. The number of previous day rainfall inputs affecting the runoff was determined by hit and trial method and it was observed that the models performed better when the number of rainfall inputs were increased from one to five i.e. from p(t-1) to p(t-1)......p(t-5).The total number of input parameters vary from four to eight when the number of previous day rainfall inputs are varied from one to five respectively as previous three day flows are always to be provided in the input vector.

## V. REFERENCES

[1] A.Q.Dar et al (2015):"Using Artificial Neural Network for real-time flood prediction in river Jhelum, J&K, India". Elixir Civil Engg. 82(2015) 32710-32713
[2] Beven, Keith J. 2001 Rainfall-runoff modelling: the primer Wiley.
[3] Dawson, C. W. and Wilby, R. L., Hydrological modelling using artificial neural networks, Progress in Physical Geography, 25(1), 80-108, 2001.
[4] Dibike, Y. B. Solomatine, D. P. 2000 “River flow forecasting using artificial neural networks” Physics and Chemistry of the Earth (B), Vol. 26, No. 1, pp. 1-7, Elsevier Science B.V.
[5] French, M. N. Krajewski, W. F. Cuykendal, R. R. 1992 “Rainfall forecasting in space and time using a neural network” Journal of Hydrology, Vol. 137, pp. 1-37, Elsevier Science B.V.
[6] Govindaraju, Rao S. 2000 “Artificial neural networks in hydrology II: hydrologic applications” Journal of Hydrologic Engineering, Vol. 5, No. 2, pp. 124-137, ASCE.
[7] Hornik, K., Stinchcombe, M. B., and White, H. (1989). Multilayer feed-forward networks are universal approximators. Neural Networks, 2(5):359–366.
[8] Maier, H. R. and Dandy, G. C., Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications, Environmental Modelling & Software, 15(1), 101-124, 2000.
[9] Meher-Homji, V. M. (1971): The Climate of Srinagar and its Variability. Geog. Rev. Ind., 33, (1): 1-14.
[10] Rientjes, T. H. M. Boekelman, R. H. 2001 Hydrological models, Lecture notes CThe4431 Faculty of Civil Engineering and Geosciences - Section of Hydrology and Ecology.