J. Indian Water Resour. Soc.,
Vol. 39, No. 3, July, 2019

MODELLING OF SPRING FLOW USING ARTIFICIAL NEURAL NETWORK

S. S. Rawat[1], Sanjay Mathur[2], H. C. Sharma[2] and P. K Singh[2]

ABSTRACT

In the mountainous region, springs are the main sources of drinking water supply as well as domestic water, therefore these springs are called the lifeline of the region. During summer, spring flow reduces significantly and sometimes they dried-up and hence people face the acute shortage of water. Spring hydrology is complex in nature and lack of a database on springs restricts the application of developed hydrological models to make these critical waters resources as a sustainable source of water in the region. In the present study, an artificial neural network (ANN) model has been developed to predict the spring flow for two springs namely, Hill Campus and Fakua from Tehri Garhwal district of Uttarakhand. The daily data of precipitation (P), temperature (T), relative humidity (RH) and spring flow (Q) from the year 1999 to the year 2003 were used to model the spring flow. First, four-year data was used to train the models, whereas, remaining one-year data was used to test the accuracy of models in predicting the springflow. Seven models differ in input parameters (P, T, RH & Q) have been developed and the best fit model was selected on the basis of capability of the network to converge normalize system error (NSE). Developed models performance were evaluated by categorize all data samples into six ranges (0-10%, 11-20%, 21-30%, 31-40%, 41-50% and >50%) of corresponding percentage error with observed data. A close agreement between observed and ANN predicted spring flow in the testing period for both the springs indicates the capability of ANN in predicting the spring flow even in data scare situation. The study also reveals that variability in training dataset results in better accuracy in the testing period.

Keywords: *Artificial neural network (ANN), Mountainous springs, Springflow.*

INTRODUCTION

Springs are the main dependable source of drinking water for more than 40 million inhabitants of the Himalayan region. For this reason, villages in the region are clustered around natural water springs (Rawat et al., 2005). Springs are naturally occurring formations and hence common water source for the local community (Mahamuni and Upasani, 2011). In addition to this, these springs are the life of several rivers and there is hardly any river, which is not fed by a spring. Hence, these springs are the lifeline of several micro-habitats in the river systems. However, spring water, a prime natural resource in mountainous regions is fast deteriorating due to dramatically increased and unplanned urban development in the region since last few decades, which altered the recharge area of these springs. Commercial and urban development, including mining, cutting of roads, highways, tunnelling, etc. are spreading rapidly across the region as the population continues to grow in the vicinity of the springs. These anthropogenic activities lead to the destruction of the internal hydrological system. For example, Valdiya and Bartarya (1989) reported a 40% reduction in spring discharge over a 35-year period (1951 to 1986) in the Kumaun Himalaya region. Mahamuni and Kulkarni (2012) identified that nearly 8,000 villages from the eastern part of Indian Himalayan Region (IHR) were facing acute water shortages. A long queue of ladies and children can be witnessed attesting to the acute shortage of water of this Himalayan region.

1. *Western Himalayan Regional Centre, National Institute of Hydrology, Jammu, J & K.*
2. *G. B. Pant University of Agricultural and Technology, Pantnagar, Uttarakhand*

Email : soban.singh@gmail.com

Manuscript No. 1510

Like gauging of rivers and groundwater, there is not any agency in the country which is accountable for gauging the flow of these vital water resources of the Himalayan region. Hence, a systematic database/inventory on springs like river and groundwater could not be generated in the country and these vital water resources could not be the part of our National Water Policy (NITI Aayog, 2018). Lack of basic information even the daily flows of major springs is the major hurdle inproper assessmentof the potential of these springs to fulfil the water demand of local populace.To overcome the local water stress of the area Agarwal et al. (2012) has given the emphasis to carry out spring-by-spring long term minimum water availability and actual water demand relationship based on daily spring flow data. Minimum versus maximum flow ratios versus time lag between rainfall and spring flow could be an indicator for identifying improvement or decay of springs in a watershed.

It is evident from the literature that several empirical, conceptual and physically-based models have been developed and applied in a different aspect of hydrology. However, their application in groundwater hydrology of spring is very selective due to lack of a basic database to understand the physical processes taking place in the hydrological system beneath the undulating surface of hilly terrain (Pingale et al., 2013; Christoph & Steffen, 2010; Negi and Joshi, 2004).

In this paper, an attempt has been made to develop the most stable and efficient artificial neural network (ANN) model having the configuration of different input parameters and different neurons in various hidden layers for predicting springflow of two natural water springs from Tehri Garhwal district of Uttarakhand state of India. The main purpose of this paper is to use an artificial neural network which has the unique capability to map input on output even in data scare situation in the field of spring hydrology.

STUDY AREA

The study was conducted on two natural water springs viz., Hill Campus spring and Fakua spring located in the College of Forestry, Ranichauri, Veer Chandra Singh Garwhwali University of Horticulture and Forestry (formerly Hill Campus of Govind Ballabh Pant University of Agricultural and Technology, Pantnagr) at Ranichauri in Tehri Garhwal district of Uttarakhand at the altitudes 2050 m and 1850 m above mean sea level, respectively (Fig. 1). This area falls under outer or lesser Himalayas and area is strongly undulating and hilly. Mean annual rainfall is about 1176 mm and ranges between 4 mm (November) and 246 mm (August). Mean maximum temperature recorded at Hill Campus, Ranichauri ranged from 10.6 ^{o}C (January) to 25.5 ^{o}C (May). The mean minimum temperature varied from 1.9 ^{o}C (January) to 14.9 ^{o}C (May). The average humidity recorded at 2 PM is minimum (30.4%) during April and maximum (83%) during August. The texture of Ranichauri soils vary with the altitude as well as with the depth of soil strata. Generally, coarse sand fraction decreases with altitude. It is higher in lower, medium in middle and quite low in higher altitude. Fine sand fraction decreases while total sand increases with depth. It has shown that Hill Campus soil series has more silt content than the other soil. This silt fraction decreases with depth. The clay fraction dominants in Hill Campus soil in comparison to Fakua soil. The catchment of Hill Campus spring having dense forest mainly oak (*Quercusleucotricophora*), deodar (*Cedrusdeodara*), burans (*Rhododendron arborium*), morpankhi (*Thujaorientalius*) etc. While maximum part of Fakua spring catchment is covered by shrubs like a wild rose (*Rosa burunii*), kirmora (*Barbarisasiatica*) and rest part by some trees of chirpine (*Pinusronburghii*), surai etc.

DATA AND METHODOLOGY

Hydro-metrological data of five years (1999-2003) was collected from the observatory of Hill Campus, Ranichauri of G.B. Pant University of Agriculture and Technology, Pantnagar, Uttarakhand and analysed. Precipitation (P), temperature (T), relative humidity (RH) and spring flow (Q) data were used for ANN modelling of spring flow.

Artificial Neural Network (ANN)

A neural network may be defined as a massively parallel connected network made up of simple processing units called *neurons,* which is capable of extracting features from the environment in which it is embedded and making this information available for use. Neural Networks are also known as neurocomputers, connectionist networks, or parallel-distributedprocessors (PDPs). The power of the neural network lies in its massively parallel-distributed structure and its versatility comes from its ability to learn by example and producing reasonable outputs even for inputs not encountered during training. The Artificial Neural Network (ANN) is a black box model which is fully based on observational data and on the calibrated input-output relationship without the description of an individual

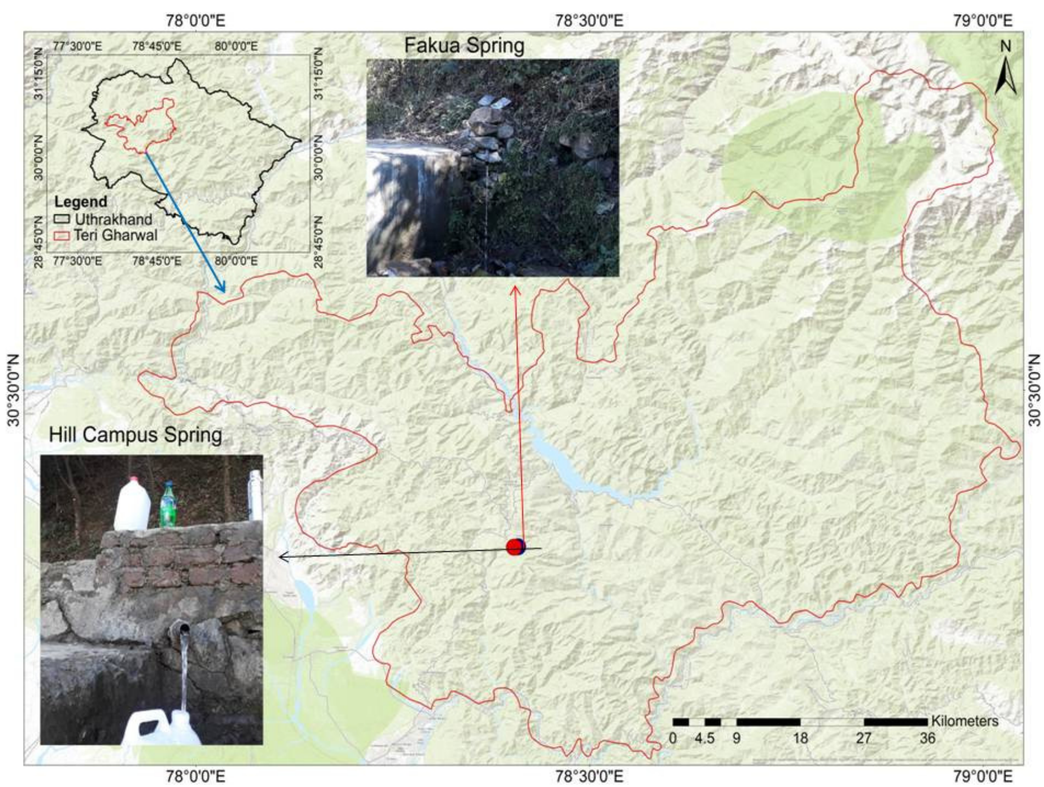

Figure 1: Location of Hill Campus and Fakua Springs in Tehri Garhwal district of Uttarakhand.

process. Neural networks learn from experience and then perform 'recognition without definition' (Kasko, 1994). Artificial neural networks are capable to handle nonlinearity of the complex systems to be modelled withflexible mathematical structure along with the activation function.

Application of Neural Networks

The application of ANN is not only in hydrology, it has been successfully applied across an extraordinary range of problem domains, in areas as diverse as medicine (Venkatesan and Anitha, 2006); aerodynamic optimization (Wei et al., 2008); construction cost forecasting (Zhigang and Yajing, 2009); Pattern recognition (Miyoung and Cheehang, 2000). As far as hydrology is concerned ANN has been applied several diverse hydrological problems and the results in each case have been very encouraging. One of the important characteristics of ANN is their adaptive nature, learning by examples (Deco and Obradovic, 1996; Haykin, 1999). ANN can find useful relationships between different inputs and outputs without attempting to reach understanding as to the nature of the phenomena. The nonlinear nature of the relationship, universal function approximation, robustness, ability to learn, and the complexity of physically based models are some of the factors that have suggested the use of ANN in hydrology (ASCE, 2000a&b).

Back-propagation artificial neural network (BPANN)

Development of back-propagation artificial neural network (BPANN) created a great impact in the field of ANN among the researchers. The back-propagation algorithm originated from Widrow and Hoff's learning rule and popularized by Rumelhart et al., (1986). It is a systematic method for training multilayer neural networks. As a result of this algorithm, multilayer perceptrons are able to solve many important practical problems.

The neural net consists of a set of the sensory input layer, one or more hidden layers of computation nodes and an output layer of computation nodes. These are also called multilayer perceptrons (MLPs) that can be trained in a supervised manner with an algorithm known as the error back-propagation algorithm (Rumelhart and McClelland, 1986). This algorithm is based on the error-correction learning rule. Figure 2 shows the architectural graph of a fully connected multilayer perceptron with two hidden layers and an output layer.

Model Development

ANN has been explicitly used in rainfall-runoff modelling, however, its use in springflow modelling is yet to be tested. In the present paper, an attempt has been made to extend the ANN methodology to simulate and predict the spring flow. The networks were trained and tested by using the data that represent different characteristics of spring's catchment and rainfall patterns. The selection of training data that represent characteristics of a catchment and metrological patterns is extremely important in modelling (Yapo*et al.,* 1996). The training data should be large enough to contain the characteristics of the catchment and to satisfy the requirements of the ANN architecture. If the information

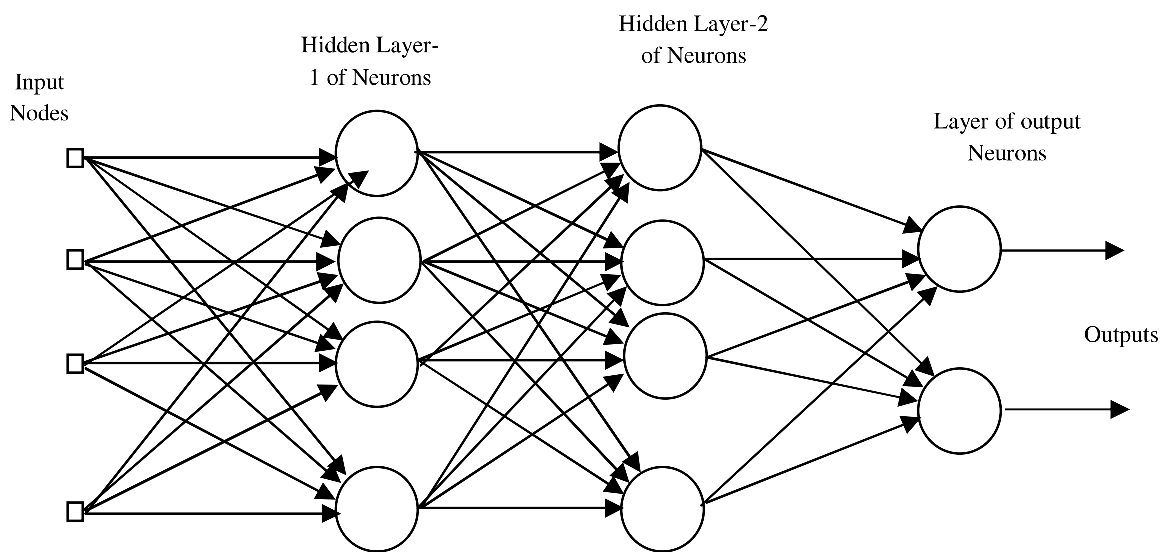

Figure 2: Architectural graph of a multilayer perception having two hidden layers

included in the training data set is insufficient it may result in increasing complexity of the network (i.e. an increase in the number of neuron or layer in a network) and will not enable the network to generalize patterns in the physical phenomena.

In this study, five years of spring data (1999-2003) were used to train and test the network. First, four years data were used to train the network and remaining one-year data was used to test the network performance, without directly including the land use characteristics of the spring's catchment. Network with various number of neurons in one or two hidden layers (i.e. network with a neuron varying from one to as many as 20) was trained for various combinations of daily precipitation (P), temperature (T), relative humidity (RH) and spring discharge (Q) in order to illustrate the effect of input variables, characteristics contained by the data in the training set on the model prediction accuracy. By trial and error method the most appropriate value for learning rate and momentum was found to be 0.5 and the number of iterations was fixed as 5000. The procedure used to model the spring flow is summarized in the following steps.

1. A simple model was selected by representing spring flow at the present time 't' i.e. Q(t) = f{p(t)}. Various ANN configurations were trained and tested by using this model (i.e. network with a number of neurons varying from one to as many as 20 in a single layer and after that go to two hidden layer network). The network having least Normalized System Error (NSE) in training and testing was then finally selected.
2. The precipitation at the time 't-1' was added as an additional input variable to the model at step 1. Now, spring flow was expressed as a function of precipitation at time 't' and 't-1' i.e., Q(t) = F{P(t), P(t-1)}. Now the network has the least NSE in training and testing was compared with those that for the best fit model at the previous step. If NSE significantly increased from the previous step, then precipitation at time 't-2' was added as another input variable to the present model i.e. Q(t) = F {P(t), P(t-1), P(t-2)}. The procedure was repeated by adding precipitation at previous time periods as an input variable until there was no significant improvement in model training and testing accuracy based on NSE.
3. After completing step 2, other input variables, such as temperature, relative humidity and spring flow of previous time periods, were added to the best fit model selected from step 2. The procedure at step 2 was repeated for each of these variables until the least NSE is achieved.

Thus, seven daily spring flow models were developed based on the procedure outlined in step 1 to 3 and the mathematical forms are presented below:

i. Model 1: Q(t) = F {P(t)}
ii. Model 2: Q(t) = F {P(t), P(t-1)}
iii. Model 3: Q(t) = F {P(t), P(t-1), P(t-2)}
iv. Model 4: Q(t) = F {P(t), P(t-1), T(t)}
v. Model 5: Q(t) = F {P(t), P(t-1), T(t), T(t-1)}
vi. Model 6: Q(t) = F {P(t), P(t-1), T(t), RH(t)}
vii. Model 7: Q(t) = F {P(t), P(t-1), T(t), Q(t-1)}

A programme in 'C' language was written and used to realize the ANN as software. After training and testing of different models listed above these models were studied on the basis of number of samples predicted by ANN in different range of percentage errors (i.e. 0-10%, 11-20%, 21-30%, 31-40%, 41-50% and >50 %).

RESULTS AND DISCUSSION

In the present study, three-layer back propagation artificial neural network (BPANN) models (different in input parameter) developed on a daily time basis for prediction of spring flow for Hill Campus and Fakua springs of Uttaranchal State. All developed models were tested for convergence of normalized system error (NSE) for 5000 iterations.

As shown in Table 1 & 2, NSE for the Model 2 for the best ANN structure (2-4-2-1) is lower than the best architecture of Model 1 (1-4-2-1). Therefore, the addition of P at (t-1) improves the accuracy of the ANN model. However, the addition of P at (t-2) did not improve the performance of the model, therefore, NSE value of Model 3 increases. This means P at (t-2) does not provide any significant information; it only increases the complexity of the network. The performance of the best fit network for Model 4, significantly improved as compared with Model 3 and 2. This shows that temperature at the time (t) provides extra information in addition to Model 2, to simulate the spring flow. Model 4 also has a simple architecture of (3-4-1) implying network becomes simpler after addition of temperature at (t). The temperature at (t-1) did not improve the model accuracy. The performance of Model 6 is slightly improved over Model 4 and Model 5. This shows that the model performs slightly better with the addition of relative humidity (RH) at the time (t). Since relative humidity (RH) is the function of temperature, the information provided by the addition of RH has already beenprovided by the factor of temperature at (t) in Model 4.

Table 1: Architecture of BPANN models for Hill Campus spring.

Model No	No. of nodes in the input layer	No. of nodes in hidden layers		No. of nodes in the output layer	Normalized system error (NSE) (5000 iterations)
		1st	2nd		
Model 1	1	4	2	1	0.002060
Model 2	2	4	2	1	0.002006
Model 3	3	4	0	1	0.002128
Model 4	3	4	0	1	0.001481
Model 5	4	4	2	1	0.001646
Model 6	4	4	2	1	0.001465
Model 7	4	8	0	1	0.000189

Table 2: Architecture of BPNN models for Fakua spring.

Model No	No. of nodes in the input layer	No. of nodes in hidden layers		No. of nodes in the output layer	Normalized system error (5000 iterations)
		1st	2nd		
Model 1	1	4	2	1	0.007214
Model 2	2	4	2	1	0.007179
Model 3	3	4	2	1	0.007295
Model 4	3	4	0	1	0.007148
Model 5	4	4	2	1	0.007423
Model 6	4	4	2	1	0.007123
Model 7	4	8	0	1	0.00080

It is evident from Tables 1 & 2 that Model 7 has quite low values of NSE as compare to other models for the same number of iterations. It clearly indicates that the performance of Model 7 is superior to other developed models. Furthermore, the addition of previous day discharge reduces the complexity of the network which shown as the simple network structure of BPANN model i.e. (4-8-1).

Beside NSE criteria, all the data were categorized in their percentage error in the prediction of spring flow of Hill Campus and Fakua springs by different models in six ranges (0-10%, 11-20%, 21-30%, 31-40%, 41-50%, >50%) and depicted in the Figures 3-6 during training and testing for Hill Campus and Fakua springs. As shown from bar diagram in Figure 3 and 4 for Hill Campus spring and Fakua spring, respectively, Models 1, 2 and 3 having more number of samples in the lower range of percentage error in comparison to Models 4, 5 and 6 in the same range for training while NSE values of models are decreasing from Models 1 to 6. It may be because of the data of the year 2001 in training set has variability ranges from 0 to 22.26 % for Hill Campus spring and 2.68 to 71.09 % for Fakua. While the average variability of the training data set is 2.1

to 98 % for Hill Campus spring and 9 to 263 %for Fakua spring. ANN Models 1, 2 and 3 are unable to predict the

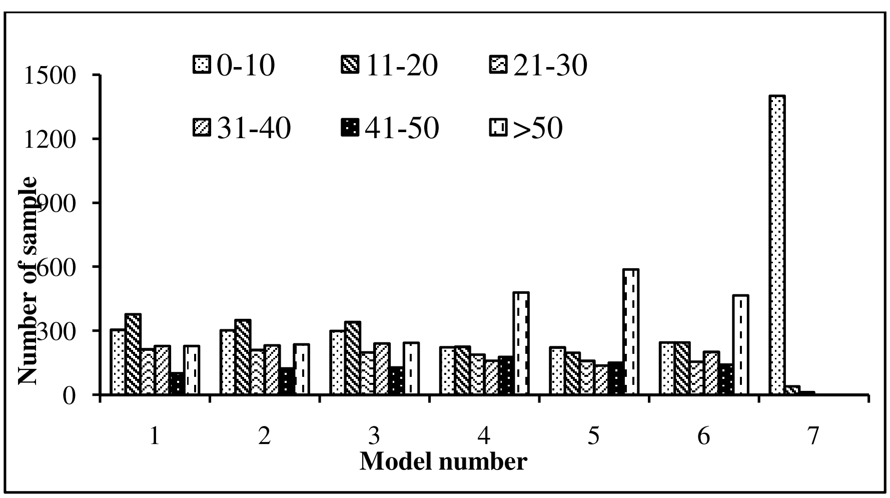

Figure 3: Bar diagram plot for the number of samples in different ranges of percentage error for Hill Campus spring predicted by different BPANN models in training

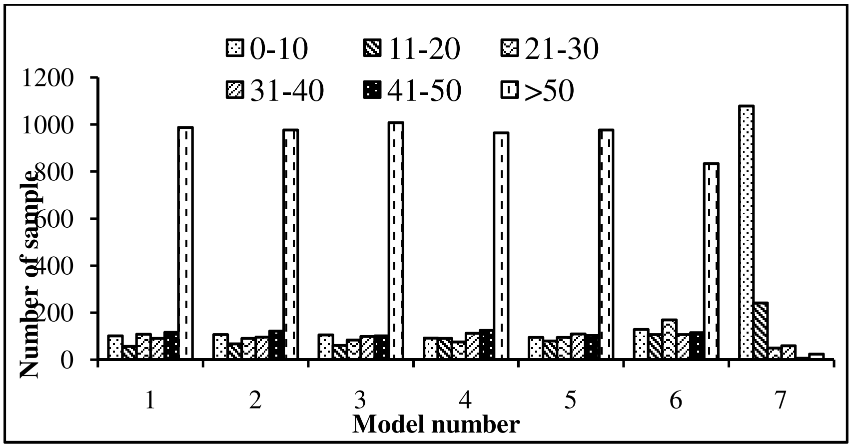

Figure 4: Bar diagram plot for the number of samples in different ranges of percentage error for Fakua spring predicted by different BPANN models in training

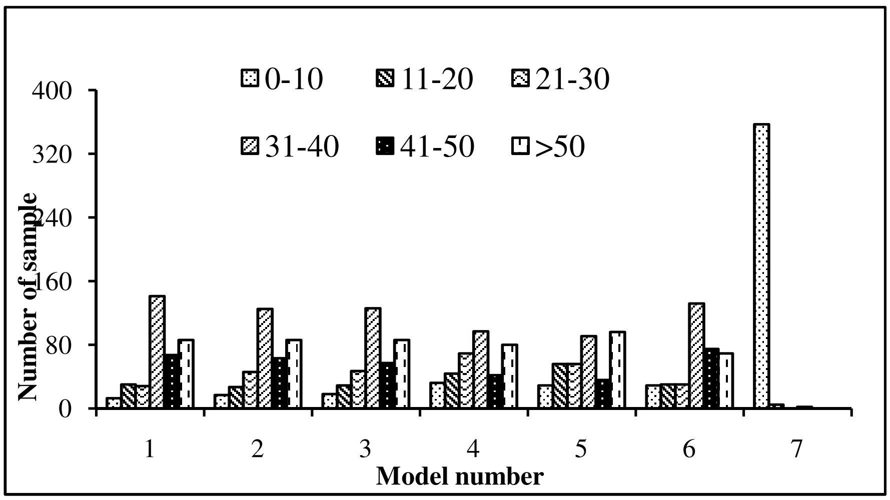

Figure 5: Bar diagram plot for the number of samples in different ranges of percentage error of Hill Campus spring predicted by different BPANN models in testing

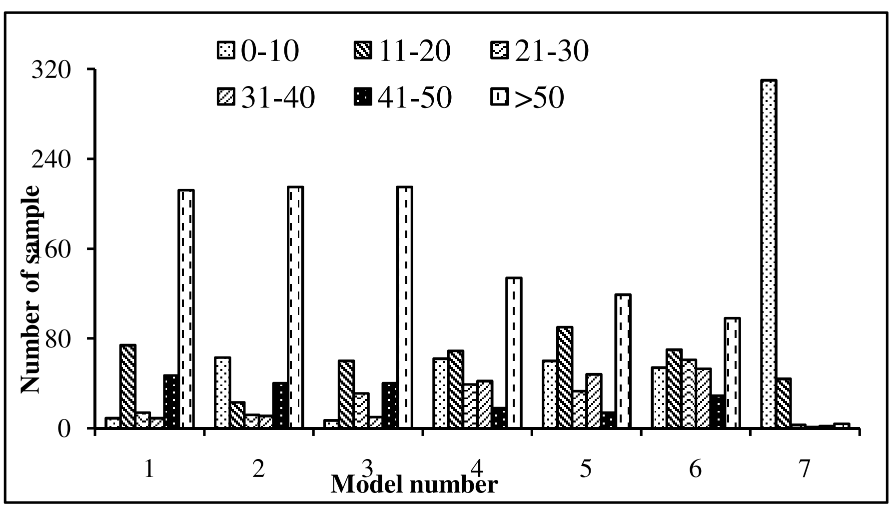

Figure 6: Bar diagram plot for the number of samples in different ranges of percentage error of Fakua spring predicted by different BPANN models in testing

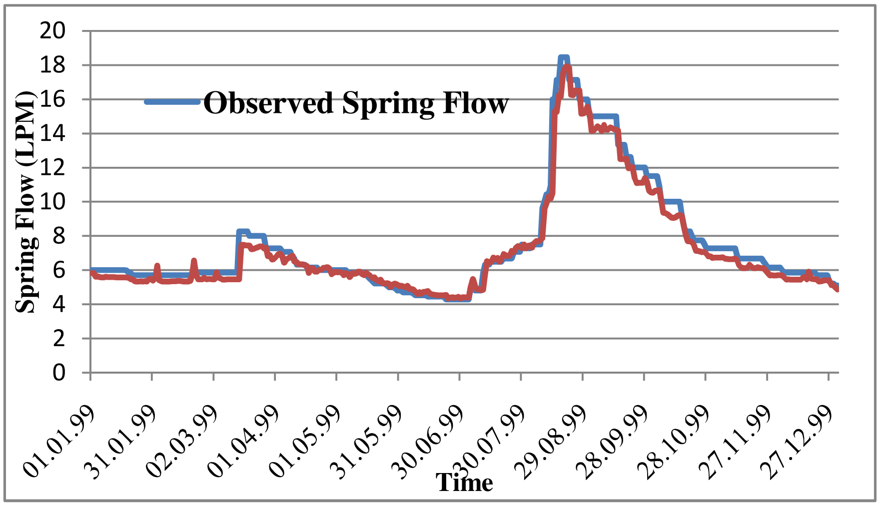

Figure 7: Observed and estimated spring flow for Hill Campus spring by ANN model during the testing period (the Year 2003) for Model 7 (4-8-1 network).

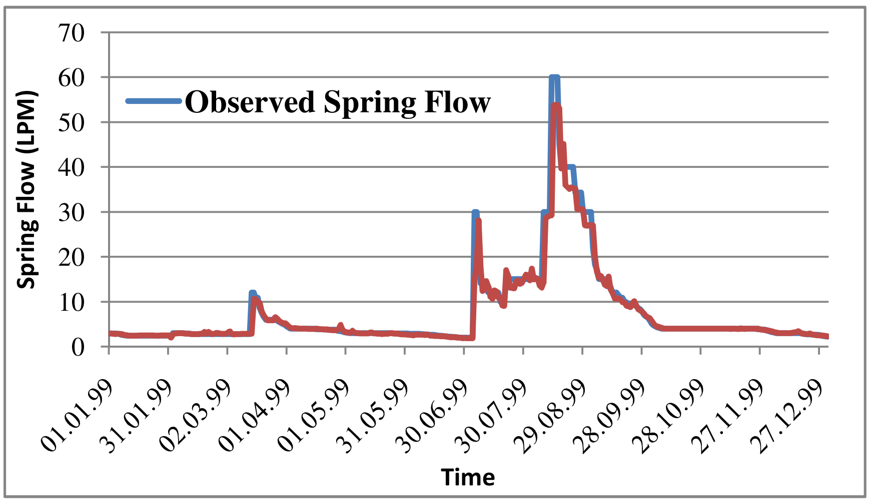

Figure 8: Observed and estimated spring flow for Fakua spring by ANN model during the testing period (the Year 2003) for Model 7 (4-8-1 network).

peak of spring hydrograph. Due to this these models (1 to 3) have number of samples in lower error range, even though they have greater NSE values in comparison to Models 4, 5 and 6 during training. But testing data set have variability ranging from 2.43% to 83.42% for Hill Campus and 9.0% to 212.12% for Fakua spring, so a number of samples increases in the lower range of percentage error with decreasing of NSE. It can be shown in Fig 5&6, most of the data predicted by Models 4,5&6 are in a lower range of error as compare to Models 1,2&3. It can be seen from Fig. 3-6, most of the data predicted by Model 7 are in 0-10% error range, however, in case of other models, most of the data have been predicted in a higher range of errors. Therefore, Model 7 [(P(t), P(t-1), T(t), Q(t-1)] is superior toother models in the prediction of spring flow.

The Scatter plot between observed spring flow and ANN predicted spring flow by Model 7during the testing period were plotted and depicted in Figs. 7&8 for Hill Campus and Fakua springs, respectively. Close agreement between observed and ANN predicted spring flows (Figs. 7&8) clearly indicates the accuracy of the ANN model in the prediction of spring flow of Hill Campus spring and Fakua spring. Fakua is highly variable spring flow (2 LPM to 60 LPM) but even this ANN is capable to predict spring flow very accurately even better than Hill Campus spring, low variability and low flow (4.1 LPM to 18.5 LPM). It also supports that training of ANN model with high variability data set resulted in better accuracy in testing.

The study reveals that Model 7 (precipitation at t, precipitation at t-1, the temperature at t and spring flow at t-1) is the best model among all models in prediction of spring flow. Unlike a fuzzy logic ruled based model (FLRBM) as developed by Pingale et al (2013), present model has capability to forecast the spring flow at daily time scale with high accuracy under limited climatological parameters. The developed model also waive-off the requirement of geomorphological parameters of the springshed which involves rigorous field work and required very fine scale springshed mapping. Developed BPANN Model can be extended to assess the effect of climate change on the spring flow using new set of climatic parameters derived from General Circulation Models (GCMs) as input of the model.

CONCLUSION

Back Propagation Artificial Neural Network (BPNN) models for spring flow have been developed by using precipitation (P), temperature (T), relative humidity (RH) and spring flow (Q) as input parameters. Springflow was successfully predicted by BPANN models without directly including the geomorphological characteristics of springshed. The effect of different input parameters has been studied in steps. This study has emphasized more on the relative study of different models having a different combination of input parameters. Therefore, learning rate (0.5), momentum rate (0.5) and the number of iteration (5000) were kept fixed and these values were decided by trial and error method.On the basis of the above discussions, the following conclusions may be drawn:

1. Precipitation is the main driving source of spring but temperature cannot be neglected in ANN modelling of spring flow;
2. Without the inclusion of the previous spring flow, high accuracy cannot be achieved in ANN spring flow modelling, and;
3. Variability in training data set leads to a better performance of the ANN model in testing.

ACKNOWLEDGMENT

Authors are grateful to the Dean, Hill Campus, Ranichauri, G. B. Pant University of Agriculture & Technology, Pantnagar for providing the data and support during the study. Also thankful to Er. Padam Singh, Assistant Professor, Agricultural Engineering, College of Forestry, Ranichauri for sharing springshed information. The comments and suggestions of the anonymous reviewers are highly appreciated.

REFERENCES

1. ASCE, (2000a). Task committee on application of artificial neural networks in hydrology. Artificial neural networks in hydrology. 1: Preliminary concepts. Journal of Hydrologic Engineering, 5(2): 115-123.
2. ASCE, (2000b). Task committee on application of artificial neural networks in hydrology. Artificial neural networks in hydrology. 11: Hydrology applications. Journal of Hydrologic Engineering, 5(2): 124-137.
3. Christoph R., Steffen B. (2010). Hydrogeological Characterisation and Modelling of Spring Catchments in a Changing Environment. Austrian Journal of Earth Sciences Volume: 103/2, PP: 106-117
4. Deco G., Obradovic D. (1996). An information theory based learning paradigm for linear feature extraction, Neuro computing, 12, 203-221.
5. NITI Aayog (2018). Report of working group I: Inventory and revival of springs in the Himalayas for water security,
6. Haykin S; (1999). Neural networks: a comprehensive foundation. New Jersey: Prentice Hall.
7. Kasko, B., (1994). Neural networks and fuzzy systems. Prentice-Hall *Inc.*, Englewood Cliffs, New Zealand.
8. Mahamuni, K., Kulkarni, H. (2012). Groundwater resources and spring hydrogeology in South Sikkim, with special reference to climate change. In: Climate change in Sikkim - Patterns, impacts and initiatives, eds., Arrawatia, M.L.; Tambe, S. Gangtok, India: Information and Public Relations Department, Government of Sikkim. Pp. 261-274.
9. Negi GCS, Joshi V. (2004). Rainfall and spring discharge patterns in two small drainage catchments in the Western Himalayan Mountains, India. Environmentalist 24:19–28.
10. Pingale S.M., Chandra, H. Sharma, H.C. Singh R. (2013). Fuzzy logic rule based modelling of natural spring flow in a hilly catchment of Tehri-Garhwal district, Uttarakhand, India. International Journal of Hydrology Science and Technology, Indersciences, 3(4):289-307.doi:10.1504/IJHST.2013.060333.
11. Rawat, S. S., Sharma, H. C., Bhar, A. K., Singh, P. K. (2005). Hydrological study of springs of Uttaranchal for sustainable development. In: International Conference on Recent advances in water resources development and management, Vol. (1), pp: 603-613.
12. Rumelhart D. E., Hinton G. E., Williams R. J. (1986). Learning internal representation by back-propagating errors. In Parallel Distributed Processing: Explorations in the Microstructure of Cognition Press, Rumelhart, D.E., McCleland, J.L. and the PDP Research Group, eds. MA: MIT.
13. Valdiya, K.S., Bartarya, S.K. (1989). Diminishing discharges of mountain springs in a part of Kumaun Himalaya. Current Science, 58(8):417-426.
14. Wei S., Zhenghong G., Yingtao Z. (2008). Application of RBF neural network ensemble to aerodynamic optimization. 46th AIAA Aerospace Sciences Meeting and Exhibit, Nevada.
15. Venkatesan P., Anitha S (2006). Application of a radial basis function neural network for diagnosis of diabetes mellitus. Current Science, Vol 91(9).
16. Yapo, P.O., Gupta, V.H. Sorooshian (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. Journal of Hydrology, 181: 23-48.
17. Zhigang J., Yajing L. (2009). The Application of RBF Neural Network on Construction Cost Forecasting. Second International Workshop on Knowledge Discovery and Data Mining.