THIRD

INDIA WATER WEEK 2015

WATER MANAGEMENT FOR SUSTAINABLE DEVELOPMENT

13-17 January, 2015

PROCEEDINGS

Organized by

Government of India
Ministry of Water Resources, River Development
and Ganga Rejuvenation

Artificial Neural Networks Modeling for Suspended sediment yield estimation over Kshipra Catchment, Madhya Pradesh

Pramod Kumar Meena
PhD Research scholar
Department of Water Resources Development and Management, Indian Institute of Technology, Roorkee, India

Deepak Khare
Professor and Head
Department of Water Resources Development and Management, Indian Institute of Technology, Roorkee, India

Manish Kumar Nema
Scientist National Institute of Hydrology,
Roorkee, India

Keywords

ANN, Suspended Sediment Yield, Algorithms, Modeling, Daily Basis

Synopsis

Artificial Neural Network is a vigorous technique to develop massive relationship between the input and output variables, and able to remove complex behavior between the water resources variables such as river sediment and discharge. AAN were developed, to predict sediment yield on a daily basis for monsoon period. Model performance has been evaluated in terms of Correlation coefficient (R), Mean squared error (MSE), Root mean squared error Ratio (RMSR) and Nash–Sutcliffe model efficiency (NSE). The basic ANN architecture was optimized in term of training algorithm, number of neurons in the hidden layer, input variables for training of the model. Twelve algorithms for training the neural network have been evaluated. Performance of the model was evaluated with number of neurons varied from 1 to 25 in the hidden layer. It was observed that predicted sediment yield better correlated to observed sediment yield (R=0. 9933 and 0.9567).

Artificial Neural Networks Modeling for Suspended sediment yield estimation over Kshipra Catchment, Madhya Pradesh

Pramod Kumar Meena
Department of Water Resources Development and Management, Indian Institute of Technology, Roorkee, *India*
pramodcae@gmail.com

Deepak Khare
Department of Water Resources Development and Management, Indian Institute of Technology, Roorkee

M. K. Nema
National Institute of Hydrology, Roorkee, India

KEYWORDS
ANN, Suspended Sediment Yield, Algorithms, Modeling, Daily Basis

ABSTRACT:
Artificial Neural Network is a vigorous technique to develop massive relationship between the input and output variables, and able to remove complex behavior between the water resources variables such as river sediment and discharge. AAN were developed, to predict sediment yield on a daily basis for monsoon period. Model performance has been evaluated in terms of Correlation coefficient (R), Mean squared error (MSE), Root mean squared error Ratio (RMSR) and Nash–Sutcliffe model efficiency (NSE). The basic ANN architecture was optimized in term of training algorithm, number of neurons in the hidden layer, input variables for training of the model. Twelve algorithms for training the neural network have been evaluated. Performance of the model was evaluated with number of neurons varied from 1 to 25 in the hidden layer. It was observed that predicted sediment yield better correlated to observed sediment yield (R=0. 9933 and 0.9567).

1. Introduction

The prediction of sediment load carried by a river, which influences hydraulic structures, is significant for many studies on river engineering, dam engineering and water management (Ayteek and Kisi et al., 2012; Singh et al., 2013). About 53 % of the total Indian geographical area suffers from deleterious effect of soil erosion and other forms of land degradation (Reddy, 1999). The involvement of many often interrelated physiographic and climatic factors makes the rainfall-sediment yield process not only very complex to understand but also extremely difficult to simulate (Zhang and Govindaraju, 2003). The evaluation of sediment yield in rivers, where total load consists predominantly of suspended load, plays an important role in the design of soil conservation and pollution control practices as well as in the design and management of dams, canals and other hydraulic structures (Caroni et. al.,1984).

Since many decades, a number of black-box, conceptual, and physically-based models have been developed for understanding and analyzing the rainfall-runoff-sediment yield process in a watershed system. The physically based classical models involves an number of catchment parameters such as topographic data, soil map, land use map, urban activities, soil moisture variation, surface roughness, etc. The spatial and temporal distribution of these parameters is significant and very complex to monitor. Therefore, these models have been developed and evaluated by simplifying important parameters and boundary conditions (Nagy et al., 2002). This may lead to less accuracy in the sediment yield computation. On the other hand, the models based on soft computing approaches, which perform better than the other conventional black box models, without going into the details of catchments characteristics have been the better alternative for the water resources engineer and planner to overcome this situation. These models are data driven and does not need the spatial data set. Cigizoglu, (2002) found that the ANN could model hysteresis in the sediment concentration-water discharge relationship which was impossible with the sediment rating curves. Tayfur et al., (2003) concluded that the ANN models predicted the suspended sediment concentration better than the physics-based models. Kisi, (2004) investigated the suspended sediment concentration using multi-layer perceptrons (MLP) and concluded that the MLP performed better than the conventional statistical method, multiple linear regression. Kisi, (2005) concluded that the ANN approach gave better results than the regression based sediment rating curve and multiple linear regression models. Cigizoglu and Kisi, (2006) developed the range-dependent neural network (RDNN) for the estimation of suspended sediments and found that that the proposed method produced better results than situations in which only a single network is trained on the entire data set to estimate suspended sediment. Raghuwanshi et al., (2006) proposed ANN models to predict both runoff and sediment yield for a small agricultural watershed, and reported that the ANN models performed better than the regression models.

In this connection it would like to mention that different models are classified based on their comprehensiveness in representing involved physical processes. With increasing comprehensiveness the models are classified as black-box, conceptual and physically based distributed models. The last of the three can be considered as best choice in a theoretical sense (Meena et al., 2014). Furthermore, the recent studies on artificial neural network applications in the area of water resource include rainfall runoff modeling (Alimohammadlou et al, 2014; Nayak et al., 2013; Shabani and Shabani, 2012; Abrahart and See, 2007; Smith and Eli, 1995), Stream flow measurement (Karran et al., 2014; Aggarwal et al., 2012; Kumar et al., 2004; Zealand et al. 1999), Sediment modeling (Kisi, 2014; Liu et al., 2013; Mustafa et al., 2011). Despite of this, on recent advancement, the neural network applications still in its infancy, as in literature the study is so far that could cover its application for suspended sediment using different algorithms on daily basis.

The Artificial Neural Networks (ANNs) based algorithms getting more popularized in these days due to flexibility of use of input variables without taking into account the basic physical mechanisms involved and it does not need to be represented into mathematical expressions with actual input variables as its components. To establish the true merits of Ann’s relative with the following specific objective: To develop an Artificial Neural Network model for sediment yield estimation using different algorithm on daily basis.

2. **Study area and data**

The present study has been carried out for the Kshipra basin which is located on the Malwa plateau in Western Madhya Pradesh (India) at an average altitude of 553 m above sea level (Fig.1). Kshipra river basin is a southern tributary of Yamuna river basin which is second largest river basin of India. The studied basin has a catchment area of 5608 km^2. It is seasonal river basin which has plenty of water during the monsoon months (June- October), but the discharge goes on decreasing after monsoon and reduces to a trickle during the summers. Over the years the river has lost its perennial nature and now runs dry for a period of 5 to 6 months per year (NWM 2011). The water of the Kshipra river is used for drinking, industrial and irrigation purposes.

2.1 Hydrology data

Observed Daily rainfall and discharge data of Four Site Dewas, Ujjain Indore and Mahidpur from period of January 1994 to December 2010), and suspended daily sediment of Mahidpur site from period of January 1994 to December 2010 were obtained from the regional office of central Water commission Jaipur India. With the help of thiessen polygon method average rainfall calculated from the kshipra river basin.

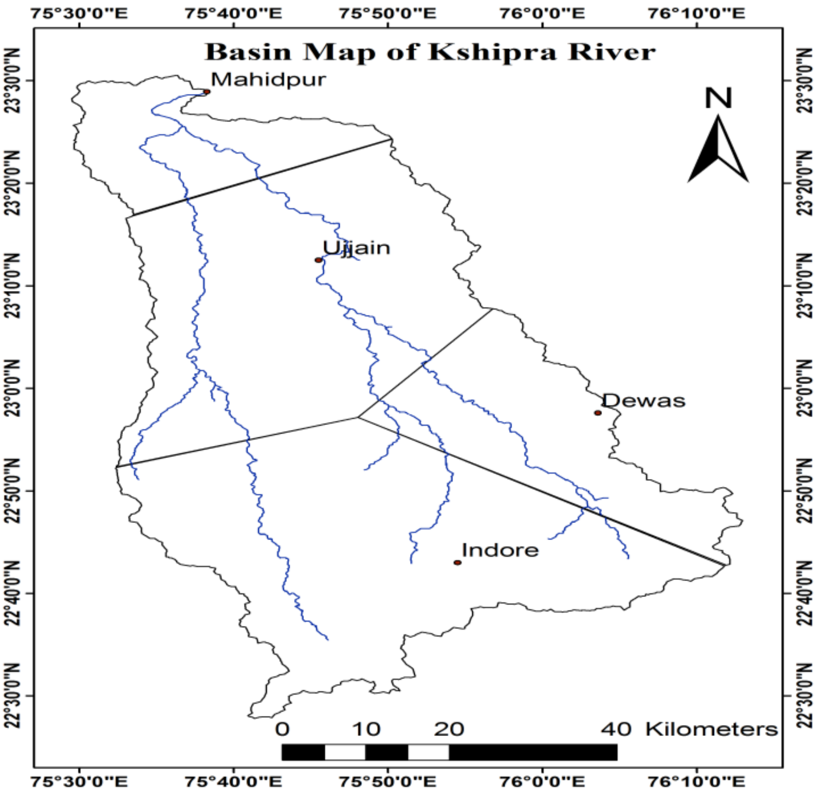

Figure1: Location of study area kshipra River Basin

3. Material and Method

3.1 ANN Model Set up

As per the literature ANN model is developed to find out those parameters whose relationship with corresponding inputs are not well defined. The model is trained under various training algorithms with number of neurons in its hidden layers and various data sets as input variables fixing different training parameters like goal, maximum number of epochs, learning rate, weight and bias matrix, layer weight and neural architecture. The motto of model development is to create a neural network to predict the spring discharge with high accuracy based on analysis of model assessment parameters. The model is often called as black box model due to its capability to do high number of iterations which cannot done by human brain and it does not follow the ordinary procedures that can be easily visualized. The model is selected by trial and error method. ANN model architecture is single layer feed forward network, which is the most commonly used neural network for the prediction of the nonlinear process. The number of the hidden layer is one. The transfer function from input to hidden layer is Tan-Sigmoid Transfer Function (Tansig) and from hidden layer to output layer is Linear Transfer function (Purelin). The Back propagation training function has been selected, which is the most common and accurate as reported by many workers. The performance function for training and testing the networks used are MSE (Mean Squared Error). The various combinations of hidden nodes and training function is done to arrive at optimum combinations to give less error. The network iterations (Epochs) were kept at 500.

The 2088 dataset is divided into training and validation data (1253 training and 835 validations) for single hidden layer in three layer feed forward back propagation algorithm. Different feed forward back propagation algorithms were applied using different transfer function and best transfer function (logsig) is chosen on the basis of performance of model. There is limited number of input parameters, so previous discharge data along-with rainfall data is taken as input variables to increase the number of input variables. The supervised learning of model is done using sediment yield (S_t) as output and discharge, Rainfall and cumulative sediment, discharge data taken as inputs as mentioned below

S_t= f (Q) ...3.1

S_t= f (Q, R) …3.2

S_t= f (Q, R, QC_2) …3.3

S_t= f (Q, R, QC_2, RC_2) … 3.4

S_t= f (Q, R, QC_2, RC_2, QC_3) … 3.5

S_t= f (Q, R, QC_2, RC_2, QC_3, RC_3) … 3.6

Where, Q- discharges, R-Avg. rainfall, RC_N- N^{th} day cumulative avg. rainfall, QC_N- N^{th} day cumulative discharge.

First, Q_t was taken as input and S_t taken as output, training and validation results obtained were not good based on model evaluation criteria NSE, RMSE, R^2 , The input is increased to 3(R_t, RC_2,RC_3) taking cumulative up to three day , same processed adopted for discharge data and performance is evaluated again.

3.2. Processing of data

The common data obtained from sources have biasness within themselves that cannot be compatible to give good results. So, processing of data is done to obtain good results and high efficiency of model.

3.3 .1 Normalization

The normalization is done because of use of log sigmoid function having limits from 0 to 1 at processing neurons. Normalization of data is done to bring dataset of different scales to a common scale by adjusting the values undergoing different mathematical and statistical processing. The motto of normalization is to produce common scale for easy comparison, processing and analysis to avoid the influence of multi-scale error.

3.3 .2 Standardization

The standardization of data is done to convert the dataset in between the range of 0 to 1. The weights and biases associated with the inputs clearly signify the dependency.

3. 3 .3 scaling of data

The scaling of data put inputs variables within a range which helps in using dataset with conditional operations. It may be simply referred as normalization. The digital computers are based on binary system, so for smooth functioning of the model the processing of data is must. The motto of scaling of data is to convert the value of inputs and outputs in the range of -1 to 1 to suit different training algorithms during calculation and to minimize the deflection of computed values from observed values.

3.4 **Performance Indicator of ANN Model**

The performance of the model can be evaluated in terms of number of criteria like mean square error (MSE), root mean square error ratio (RMSR), correlation coefficients (R), and Nash –Sutcliffe Efficiency (NSE).

4. Result and Discussion

4.1 ANN model for sediment

Initial an ANN base model has been developed with Levenverg-Marquardt training algorithm. 7 Number of neurons and single hidden layer. Six models with different predictor variables have been developed.

Table 1. Performance of Neural Network Models with Different Input Variables.

Sediment yield function	Model	R		MSE		RMSR		NSE	
		Trg	Val	Trg	Val	Trg	Val	Trg	Val
Q	Model 1	0.533	0.769	2.487	1.226	0.915	0.851	0.161	0.524
Q,R	Model 2	0.680	0.866	1.021	0.031	0.822	0.637	0.324	0.743
Q,R,QC_2	Model 3	0.896	0.893	0.421	0.034	0.529	0.699	0.720	0.740
Q,R,QC_2,RC_2	Model 4	0.716	0.842	0.557	0.075	0.736	0.960	0.457	0.485
Q,R,QC_2,RC_2,QC_3	Model 5	0.640	0.858	1.015	0.047	0.822	0.758	0.323	0.661
Q,R,QC_2,RC_2,QC_3,RC_3	Model 6	0.578	0.817	0.337	0.148	0.853	0.871	0.272	0.571

(Trg: Training, Val: Validation)

Performance of these six models has been summarized in Table 1. It can be observed that value of R during validation of the Model-3(0.893) is slightly higher Model-2 (0.866). However value R during training of the model Model-3 (0.896) is higher than other model. RMSR, MSE and NSE for the model Model-3 are higher than other model; hence model Model-3 has been selected for farther refinement.

4.2 Training Algorithms for sediment yield model

The Model-3 has been tested with the different training algorithm. For developing the ANN based sediment yield estimation, performance of 12 training algorithms were evaluated. The model Model-3 was developed using Levenverg Marquardt Algorithm (trainlm).The best training algorithm in the hidden layer of ANN model can be determined by trial and error, at which a model perform better.

Table 2 Indicates that training algorithm “traincgp” resulted model with satisfactory value of correlation coefficient as 0.889 and 0.939 during training and validation, respectively. Model performance indicator; MSE with scaled estimate and target is lowest as 0.348 and 0.028 during training and validation, respectively. RMSR has been worked out as 0.518 and 0.829; and NSE as 0.731 and 0.840 during training and validation, respectively.

Table 2. Performance of different training algorithm methods for ANN based sediment yield modeling.

Training Function Model- 3	R		MSE		RMSR		NSE	
	Trg	Val	Trg	Val	Trg	Val	Trg	Val
trainlm	0.896	0.893	0.421	0.034	0.529	0.699	0.720	0.740
traingd	0.864	0.930	0.436	0.024	0.579	0.792	0.664	0.840
trainbfg	0.917	0.848	0.113	0.042	0.481	0.933	0.769	0.697
traincgf	0.735	0.952	0.585	0.028	0.736	0.665	0.458	0.819
trainrp	0.690	0.948	1.025	0.013	0.809	0.484	0.345	0.876
trainbr	0.902	0.913	9.205	6.716	0.484	0.918	0.765	0.826
trainscg	0.706	0.955	0.850	0.036	0.753	0.775	0.432	0.709
traingda	0.757	0.830	0.921	0.061	0.761	0.773	0.420	0.684
traingdx	0.680	0.866	1.021	0.031	0.822	0.637	0.324	0.743
traincgp	0.889	0.939	0.348	0.028	0.518	0.829	0.731	0.859
traincgb	0.789	0.929	0.518	0.061	0.715	0.697	0.489	0.840
trainoss	0.852	0.901	0.586	0.055	0.678	0.652	0.541	0.811

(Trg: Training, Val: Validation)

4.3 Selection of optimum number of neurons in the hidden layer for the sediment yield

Increasing the number of neurons in the hidden layer, the network get an over fit, that is the net have problem to generalize. To determine the optimum number of neuron, at which network should have to perform its best, trial and error method is applied. Selection of optimum number of neurons is essential part sediment ANN model development. The model Model-3 with learning function “traincgp” and normalization function “mapstd” trained with 60 percent of data has been evaluated for optimum number of neurons. Neurons in the hidden layer have been varied from 1 to 25.

Table 3. Performance of neural network with different number of neurons

Neuron	R		MSE		RMSR		NSE	
	Trg	Val	Trg	Val	Trg	Val	Trg	Val
1	0.788	0.944	0.511	0.028	0.717	0.579	0.485	0.889
2	0.737	0.984	0.423	0.016	0.692	0.877	0.521	0.917
3	0.877	0.926	0.337	0.039	0.527	0.979	0.722	0.853
4	0.992	0.948	0.025	0.012	0.126	0.457	0.984	0.882
6	0.679	0.905	0.871	0.065	0.746	1.136	0.443	0.489
7	0.993	0.944	0.017	0.017	0.122	0.459	0.985	0.871
9	0.932	0.911	0.075	0.031	0.423	0.863	0.821	0.820
11	0.993	0.936	0.007	0.023	0.122	0.506	0.985	0.833
13	0.993	0.957	0.021	0.010	0.116	0.417	0.987	0.908
15	0.997	0.944	0.009	0.017	0.080	0.483	0.994	0.846
17	0.912	0.934	0.023	0.070	0.430	0.641	0.815	0.835
19	0.784	0.935	0.532	0.041	0.657	0.866	0.568	0.752
21	0.889	0.922	0.182	0.020	0.482	0.887	0.768	0.849
23	0.810	0.921	0.423	0.071	0.621	0.822	0.614	0.706
25	0.863	0.934	0.339	0.058	0.522	1.148	0.728	0.554

(Trg: Training, Val: Validation)

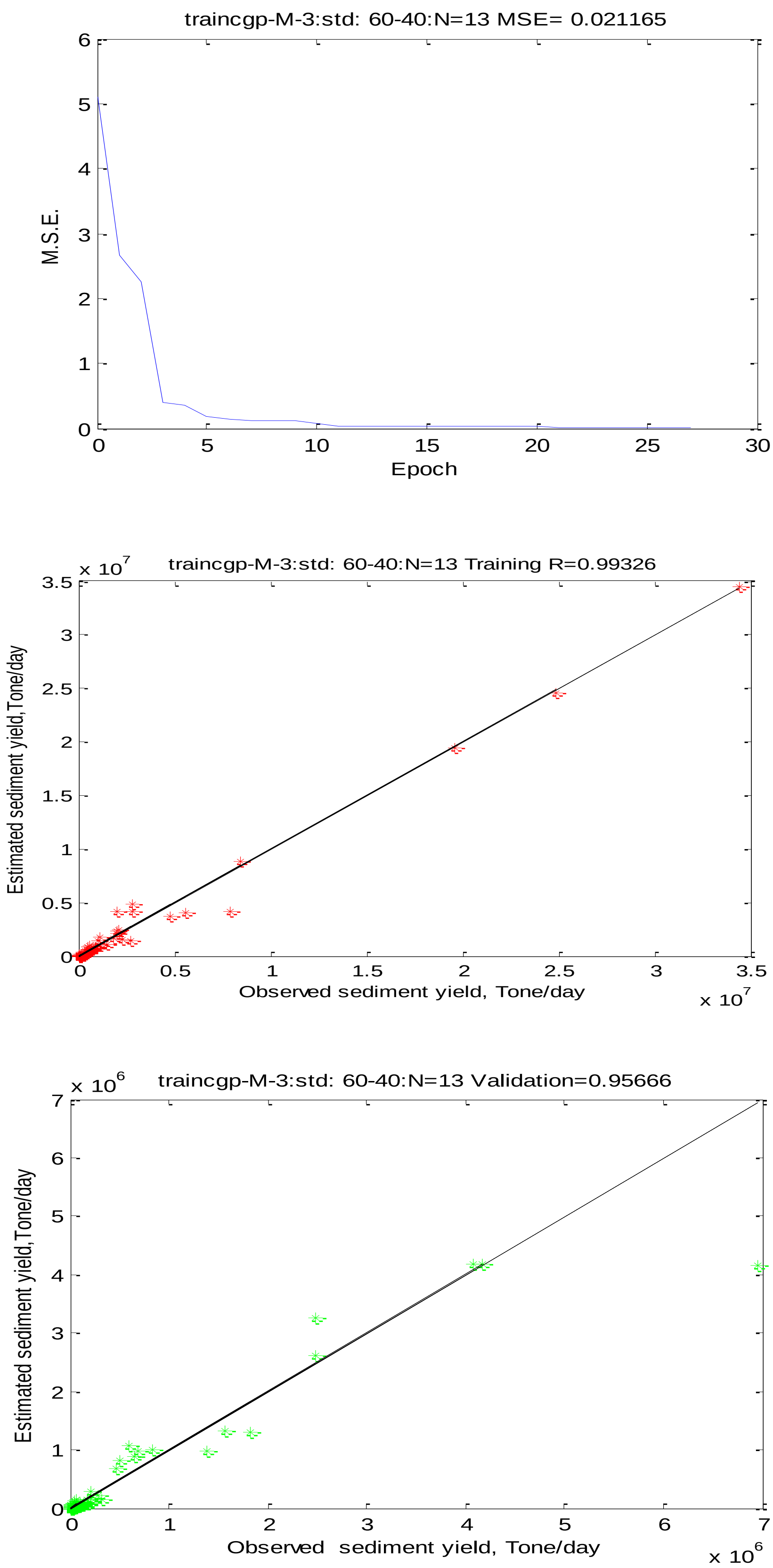

Figure 2. Performance of Model-3 model with "traincgp" training algorithm at 13 neurons for sediment yield estimation

On comparison of performance parameters presented in Table 3, it can be stated that model Model-3 trained with "traincgp" algorithm, "mapstd" normalization function and 13 neurons performed best. It's performance a shown on fig 2.

5. Conclusion

Artificial Neural Network (ANN) model for estimation of sediment yield was developed. The model has one input layer, one hidden layer and one output layer. Method normalizes the data was "std" which transform data such that mean is zero and standard deviation is unity. The input dataset for daily suspended sediment yield modeling includes daily rainfall, daily discharge and cumulative rainfall, discharge for 1 to 3 day considered as input since 1994 to 2010 on monsoon basis. ANN models were developed in Neural Network Module of MATLAB. Model performance has been evaluated in terms of R, MSE, RMSR and NSE. The basic ANN architecture was optimized in term of training algorithm, number of neurons in the hidden layer, input variables for training of the model. Twelve algorithms for training the neural network have been evaluated. Performance of the model was evaluated with number of neurons varied from 1 to 25 in the hidden layer.

From the study carried out following salient points emerged.

1: ANN model with 3 input variables is found better i.e. average rainfall, discharge and two day cumulative average Rainfall.

2: Traincgp" training algorithm of back propagation performed better than other twelve algorithms evaluated.

3: Thirteen neurons in the hidden layer of ANN model performed better.

Highest value of correlation coefficient between estimated and observed sediment yield 0.993 and 0.956 during training and validation by ANN model. The ANN sediment model with "traincgp" algorithm, 13 numbers of neurons, 60 percent and 40 percent length of record for training and validation with 3 input variables is found best model for suspended sediment yield estimation.

Acknowledgment: Authors are thankful to regional office of central water commission, Jaipur India for providing valuable data that could helpful for present study.

References

1. Abrahart, R. J. and See, L. M.(2007): Neural network emulation of a rainfall-runoff model Hydrol. Earth Syst. Sci. Discuss., 4(1), 287–326.
2. Aggarwal, S. K., Goel, A. and Singh, V. P. (2012): Stage and discharge forecasting by SVM and ANN techniques, Water Resour. Manage, 26(13), 3705–3724.
3. Alimohammadlou, Y., Najafi, A., and Gokceoglu, C. (2014): Estimation of rainfall-induced landslides using ANN and fuzzy clustering methods: A case study in Saeen Slope, Azerbaijan province, Iran. CATENA 120: 149-162.
4. Ayteek, A. and Kisi, O. (2008): A genetic programming approach to suspended sediment modeling, J Hydrol, 351:288–298.
5. Caroni, E., Singh, V.P. and Ubertini, L. (1984): Rainfall-runoff-sediment yield relation by stochastic modeling, Hydrol Sci J, 29(2):203–218.

6. Cigizoglu, H.K. (2002): Suspended sediment estimation for rivers using Artificial Neural Networks and Sediment Rating Curves, Turk J Eng Environ Sci, 26:27–36.
7. Karran, D. J., Morin, E. and Adamowski, J. (2014): Multi-step streamflow forecasting using data-driven non-linear methods in contrasting climate regimes. J. Hydroinf., 16(3), 671–689.
8. Kisi, O. (2014): Discussion of Comparison of Artificial Neural Network Models for Sediment Yield Prediction at Single Gauging Station of Watershed in Eastern India" by Ajai Singh, MohdImtiyaz, RK Isaac, and DM Denis, Journal of Hydrologic Engineering, 19(3), 661-662.
9. Kisi, O. (2005) Suspended sediment estimation using neuro-fuzzy and neural network approaches, Hydrol Sci J,50(4):683–696.
10. Kisi,O. (2004): Multi-Layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation, Hydrol Sci J, 49(6):1025–1040.
11. Kisi,O. ,Ozkan, C. and Akay, B (2012): Modeling discharge-sediment relationship using neural networks with artificial bee colony algorithm, J Hydrol, 428–429:94–103.
12. Kumar D.N., Raju S.K., and Sathish T. (2004): River flow forecasting using recurrent neural networks, Water Resources Management 18: 143-161.
13. Liu, Q. J., Shi, Z. H., Fang, N. F., Zhu, H. D. and Ai, L. (2013): Modeling the daily suspended sediment concentration in a hyper concentrated river on the Loess Plateau, China, using the Wavelet–ANN approach, Geomorphology, 186, 181-190.
14. Meena, P.K., Hardaha, M.K., Khare, D., and Mondal, A., (2014): Neural Method for Site-Specific Yield Prediction.In Proceedings of the Third International Conference on Soft Computing for Problem Solving (pp. 237-246). Springer India
15. Mustafa, M. R. U., Isa, M. H. and Bhuiyan, R. R. (2011): Prediction of river suspended sediment load using radial basis function neural network-a case study in Malaysia, In National Postgraduate Conference (NPC), 2011 (pp. 1-4). IEEE.
16. Nagy, H.M., Watanabe, K., and Hirano, M., (2002): Prediction of sediment load concentration in rivers using artificial neural networks, J Hydrol Eng, 128(6):588–595.
17. Nayak, P. C., Venkatesh, B., Krishna, B., and Jain, S. K. (2013): Rainfall runoff modeling using conceptual, data driven, and wavelet based computing approach, J. Hydrol., 493, 57–67.
18. NWM (2011): TA 7417- IND: Support for the National Action Plan on Climate Change Support to the National Water Mission. Government of India Governments of Punjab, Madhya Pradesh Tamil Nadu and Asian Development Bank
19. Raghuwanshi, N.S., Singh, R. and Reddy, L.S. (2006): Runoff and sediment yield modeling using Artificial Neural Networks: Upper Siwane River, India, J Hydrol Eng ASCE, 11(1):71–79.
20. Reddy, M.S.(1999): Theme paper on "Water: vision 2050. Indian Water Resources Soc. Roorkee, pp. 51–53.
21. Shabani, M., and Shabani, N. (2012): Application of Artificial Neural Network in Instantaneous peak flow estimation for Kharestan Watershed, Iran , Journal of Resources and Ecology,3(4) 379-383.
22. Singh, A., Imtiyaz, M., Isaac, R.K. and Denis, D.M. (2013): Comparison of artificial neural network models for sediment yield prediction at single Gauging Station of Watershed in Eastern India, J Hydrol Eng 18(1):115–120.
23. Smith, J. and Eli, R. M. (1995): Neural network models of rainfall runoff process" Journal of Water Resources Planning & Management. 121(6), 499-508
24. Tayfur, G., Ozdemir, S., Singh, V.P., (2003): Fuzzy logic algorithm for runoff-induced sediment transport from bare soil surfaces. Adv Water Resour 26:1249–1256.
25. Zealand, C. M., Burn, D. H. and Simonovic, S. P., (1999): Short term stream flow forecasting using artificial neural networks, Journal of Hydrology, 214, 32–48.
26. Zhang, B., and Govindaraju, R.S. (2003): Geomorphology-based artificial neural networks (GANNs) for estimation of direction runoff over watersheds, J Hydrol, 273:18–34.

*****

Search us : www.indiawaterweek.in

DESIGNING & PRINTING AT : BHAGWATI PRINTERS PVT. LTD., NEW DELHI TEL.: 011-45095236/37