

Optimization Of The Water Use In The River Damodar In West Bengal In India: An Integrated Multi-Reservoir System With The Help Of Artificial Neural Network

Mrinmoy Majumder, Research Scholar, School of Water Resources Engineering, Jadavpur University, Kolkata, India, mmajumder15@gmail.com

Pankaj Kumar Roy, Lecturer, School of Water Resources Engineering, Jadavpur University, Kolkata, India, pk1roy@yahoo.co.in

Asis Mazumdar, Professor & Director, School of Water Resources Engineering, Jadavpur University, Kolkata, India, asismazumdar@yahoo.com

Abstract

The present study attempts to optimize the outflow and the water distribution of entire Damodar catchment with the help of Artificial Neural Network(ANN). It is worth to mention that the large river valley project like Damodar Valley Corporation (DVC) with four dams and one barrage makes the simulation of Damodar more challenging, as it becomes a controlled flow system. Event of heavy floods are common which necessitates the development of a proper optimization of the water distribution. 8 inputs and 24 output parameters are configured and with the help of NeuroDimension 5, a commercial software. 4 ANN Models are prepared. The data from 1996 to 2000 was used to create and test the models. The model with minimum average mean square error (MSE) (2.55%) was selected and used for the optimization. The distribution of water by the model shows a priority wise selection giving the highest priority to the most important use of the local area of the adjacent reservoir.

Introduction

Neuro-morphic modeling techniques are now well established methods for describing physical processes occurring in the aquatic environment. The development in information technology over the last decade has presented opportunities of extended computational ability together with improved data manipulation, storage and retrieval. As a result neuro-morphic models are now being used more extensively in the management, design and operation of water based assets. The reason behind this is that in many years of applications pertaining to the complex flow systems, the demands on computing time are of such magnitude, which is far from acceptable.

Many new storage projects worldwide were failing to produce benefits that provided economic justification for their development. The result of this is clearly the over or under use of reservoirs, flood, drought, non-systematic supply of water. The current reservoir operational policies fail to consider integrated operations. The system-wide focus greatly multiplies alternative operational policies and further complicated by conflicting objectives, stochastic hydrology/climate change, uncertain demands etc.

DVC is the first multi purpose integrated river valley projects of independent India came into existence on 7th July 1948. In developing areas with rapidly growing population, expanding agricultural and industrial activities have, besides the impact of climatic changes, resulted in stressed conditions in the field of freshwater availability in different parts, if not the whole, of the basins. Per capita freshwater availability of water is declining at such a fast rate that it is expected that by the next 15-20 years, if not earlier, supply of fresh water would be highly stressed. West Bengal and Jharkhand, both of these states depend primarily on seasonal precipitation for their water supplies and severe water shortage during the period of drought (Roy et al, 2004).

The Damodar river basin up to the outlet of the DVC system has been given the main investigation because the water supply arrangement of both the states is considerably dependant upon that part of river basin. Thus the region up to the outlet of the DVC system is especially vulnerable to potential changes in regional temperature and precipitation pattern. Yearly 3,42,105 ha of kharif crop, 22,270 ha of rabi crop and on an average 30,000 ha under boro crop are being irrigated in the districts of Burdwan, Hooghly, Bankura and Howrah in the state of West Bengal from the water of DVC dams and regulated by the Durgapur barrage through existing canal network. Yearly about 3,45,000 acre feet of water is being released from DVC dams to cater the water demand of about 140 industries (mainly coal washeries, steel plants, municipal supply, railways, etc.) situated in the states of Jharkhand and West Bengal, India. The present study aims to create a specific pattern of the use of water so that proper use or quantity of water needed for different purposes can be optimized. The model mainly considers the quantity of water for hydropower, industry and domestic purposes in the DVC catchments.

Study Area

The Damodar river which lies between the latitudes 23°30'N and 24°19'N and longitudes 85°31'E and 87°21'E, originates from the Palamu Hills of Chota Nagpur at an elevation of about 610 m above mean sea level as shown in Fig.1. It flows in a south easterly direction, entering the deltaic plains below Raniganj in Burdwan district, of West Bengal, India. Near Burdwan the river abruptly changes its course to southerly direction and joins the Hughli river about 48 km below Kolkata. The slope during the first 241 km is about 1.89 m per km. During the next 161 km is about 0.57 m per km, while the same is about 0.19 m per km in subsequent 145 km. The river is fed by six streams of which the principal tributary Barakar joins it where river Damodar emerges from the Palamu Hills. The four main multipurpose reservoirs located at Tilaiya, Konar, Maithon, Panchet and a Barrage at Durgapur were commissioned during 1953–1959. Another tributary Khudia, whose catchment is neither intercepted by Maithon nor Panchet reservoirs, joins Damodar near its confluence with Barakar. In the plains river splits into several channels and ultimately joins the river Roopnarayan & Houghli. The total length of the river is about 541 km. The total catchment area of the river is 28,015 km², of which 26,015 km² the catchment covers the region upto Durgapur Barrage, the outlet of the DVC system. The upper portion of the catchment consists rough, hilly regions, where as the lower portion is of the nature of flat deltaic plane (Roy et al, 2004).

Among the four main reservoirs Tilaiya (4 MW), Maithon (60 MW) and Panchet (80 MW) are used to generate hydropower. Konar and other small dams are presently not used for hydropower.

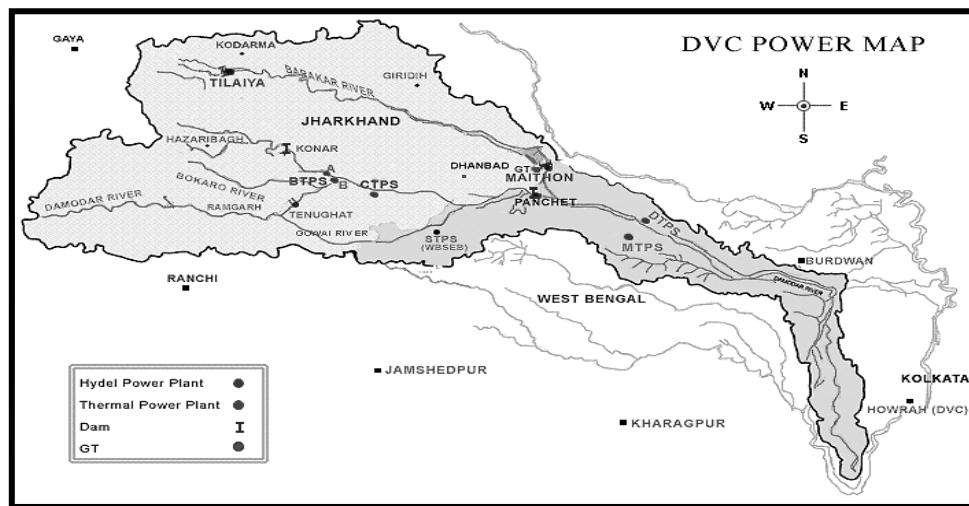


Fig.1: Damodar River Basin (Roy et al, 2005)

Methodology

The study has been done mainly for optimizing the operation of four main reservoirs namely Tilaiya, Konar, Maithon and Panchet in the DVC system. The other small reservoir like Tenughat is not considered to reduce the complexity since the influence of this reservoir is negligible. In this model, the input and output parameters were identified by using the historical hydrological data (discharge, reservoir level etc) which has been collected from DVC. The use of water such as domestic, industrial and power generation for all the four reservoirs has been calculated in the following way:

If, $Outflow = O$

Water Requirement = W , Minimum Required = W_m , available with the regulatory body.

Then requirement of water for one type of use, say H , is given as:

Criteria: $O > W_m$,

$$H = W/O * 100$$

Criteria: $O < W_m$

H = Water unavailable for the designated use.

Although the water required for generation of hydro power, at an instant, is calculated in a different way. The percentage of water required to generate hydropower at maximum and minimum head are considered.

As every artificial neural network needs to be trained and tested before it is applied for forecasting. So the total data is divided into three parts, namely training (80%), cross validation (10%) and testing (10%). Neuro Dimension 5, a commercial software is used to train and test the models. The model DVC-GFF-1(Table-2) is selected based on the best correlation and least MSE than the other three models. The graphical representation of this model is also depicted in Fig.2 and elaborated in Table-1.

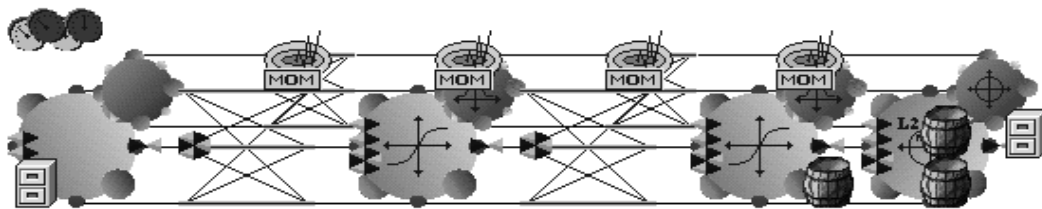
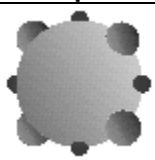
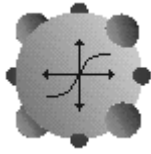


Fig 2 : The Graphical Representation of the Model(DVC-GFF-1)(Table – 1)

Table – 1		
Explanation of the graphical representation of the selected model(DVC-GFF-1)		
Graphical Component	Explanation	Equations
 <p>Axon</p>	<p>The Axon's activation function is the identity map. It is normally used just as a storage unit. Recall however, that all axons have a summing junction at their input and a node junction at their output.</p>	$f(x_i, w_i) = x_i$ <p>Where, x_i = Input w_i = Weight at ith iteration</p>



Tanh Function

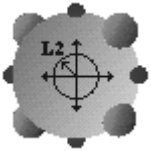
The TanhAxon applies a bias and tanh function to each neuron in the layer. This will squash the range of each neuron in the layer to between -1 and 1. Such nonlinear elements provide a network with the ability to make soft decisions.

$$f(x_i, w_i) = \tanh [x_i^{lin}]$$

Where,

$$x_i^{lin} = \beta x_i$$

x_i^{lin} is the scaled and offset activity inherited from the Linear Axon.



L2 Criterion

The L2Criterion implements the quadratic cost function. This is by far the most applied cost function in adaptive systems. The error reported to the supervised learning procedure is simply the squared Euclidean distance between the network's output and the desired response.

$$J(t) = \frac{1}{2} \sum_i (d_i(t) - y_i(t))^2$$

$$e_i(t) = -(d_i(t) - y_i(t))$$

Where,

d_i = Desired output(function of time)

y_i = Actual Output(function of time)

$J(t)$ = Cost function

$E(t)$ = Error Function



BackStaticControl

The BackStaticControl component is used in conjunction with the StaticControl component. Static backpropagation assumes that the output of a network is strictly a function of its present input (i.e., the network topology is static). In this case, the gradients and sensitivities are only dependent on the error and activations from the current time step.



StaticControl

The StaticControl implements data flow for static backpropagation. It expects a static input and a static desired response, from which an error is obtained. The error is propagated through the dual system (backprop plane).





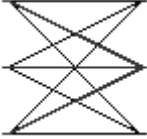
Connector

Connecting the data control from Input to Output.



Storage

The DataStorage component collects data from an access point of the component attached below

		and stores it in a circular buffer. The DataStorage component periodically sends a message to the components stacked above so that they can re-process the data or refresh their displays.
	File	The File component reads data from the computers file system.
	Momentum	<p>The Momentum provides the gradient descent with some inertia, so that it tends to move along a direction that is the average estimate for down. The amount of inertia (i.e., how much of the past to average over) is dictated by the momentum parameter, r. The higher the momentum, the more it smoothes the gradient estimate and the less effect a single change in the gradient has on the weight change.</p> <p>Synapse carries the weight updates from the inputs to hidden layer and from then to the output layer.</p>
	Synapse	

$$\Delta w_i(n+1) = \eta_i \nabla w_i + \rho \Delta w_i(n)$$

Δw_i = Weight update on n+1 th layer
 η = Momentum
n = No. of layer

Table – 2
Explanation of the deduced models

Model Name	Model Type	Inputs	Hidden	Outputs	Training Algorithm	R ² /M.S.E(%)
DVC - GFF - 1	GFF - ANN	Daily reservoir Level and Inflow of all the four dams. Total 8.	1	Daily reservoir outflow, surplus, water required for domestic, industrial and power generation uses of all the four dams. Total 24	Neural	0.94/2.5
DVC - GFF - 2	GFF - ANN	Daily reservoir Level and Inflow of all the four dams. Total 8.	1	Daily reservoir outflow, surplus, water required for domestic, industrial and power generation uses of all the four dams. Total 24	Genetic	0.51/2.46
DVC - FA	-	Daily reservoir	1	Daily reservoir	Neural	0.01/12.5

FA – 3	ANN	Level and Inflow of all the four dams. Total 8.		outflow, surplus, water required for domestic, industrial and power generation uses of all the four dams. Total 24		
DVC - FA - FA – 4	- ANN	Daily reservoir Level and Inflow of all the four dams. Total 8.	1	Daily reservoir outflow, surplus, water required for domestic, industrial and power generation uses of all the four dams. Total 24	Genetic	0.01/12.35

NOTE : GFF-ANN = Model prepared with the help of General Feed Forward algorithm, FA-ANN = Model prepared with the help of Functional Approximation Artificial Neural Network, Neural = Model prepared with the help of General Feed Forward/Functional Approximation and trained in neural algorithm, Genetic = Model prepared with the help of General Feed Forward/Functional Approximation and trained in genetic algorithm

Result And Discussion

Artificial Neural Network

An (ANN) is a flexible mathematical structure that is capable of identifying complex nonlinear relationships between input and output data sets. Neural networks provide model-free solutions. In short the modeling process starts with defining the purpose/goal of the model. Next step involves identification of key input and output variables in the system. These system variables are mapped in the ANN. Then structure of ANN is adjusted by trial and error making decisions on number of nodes in the hidden layer etc. Neural network is then run to test the behavior of the system. Network is evaluated by comparing the predicted values with observed one. Based on the evaluation, improvements are made to the network structure. When network performance is satisfactory it is ready for forecasting. In order to evaluate the performance of the neural network, percentage error and correlation were calculated for each output parameter. The deviation of actual statistics measures the percentage error in observed and predicted values. 0.0 is the best. Positive values indicate overestimation and negative values indicate underestimation.

Mean Square Error (M.S.E.)

$$\frac{\overline{(ET_i - ET_i)}}{ET_i} \times 100 \dots\dots - 1$$

The correlation statistics measures the linear correlation between observed and forecasted values; the optimal value is 1.0. The correlation calculation returns the covariance of two data sets divided by the product of their standard deviations. Correlation can be used to determine whether two ranges of data move together that is, whether large values of one set are associated with large values of the other (positive correlation), whether small values of one set are associated with large values of the other (negative correlation), or whether values in both sets are unrelated (correlation near zero).

Correlation R²

$$\frac{SS_3^2}{SS_1 SS_2} \dots\dots - 2$$

$$SS_1 = (ET_i - \overline{ET_i})^2$$

$$SS_2 = (\overline{\overline{ET_i}} - \overline{\overline{ET_i}})^2$$

$$SS_3 = (ET_i - \overline{ET_i})^2 (\overline{\overline{ET_i}} - \overline{\overline{ET_i}})^2$$

Where ET_i = Observed Outputs employed in the analysis; $\overline{ET_i}$ = Predicted Outputs; $\overline{\overline{ET_i}}$ = Arithmetic average of actual Outputs; $\overline{\overline{\overline{ET_i}}}$ = Arithmetic average of predicted Outputs

General Feed Forward

Generalized feed forward networks are a generalization of the MLP such that connections can jump over one or more layers. In theory, a MLP can solve any problem that a generalized feed forward network can solve. In practice, however, generalized feed forward networks often solve the problem much more efficiently. A classic example of this is the two spiral problem. Without describing the problem, it suffices to say that a standard MLP requires hundreds of times more training epochs than the generalized feed forward network containing the same number of processing elements (PEs).

Multi-Layer Perceptron (MLP)

MLP is one of the most widely implemented neural network topologies. The article by Lippman, 1986 is probably one of the best references for the computational capabilities of MLPs. Generally speaking, for static pattern classification, the MLP with two hidden layers is a universal pattern classifier. In other words, the discriminant functions can take any shape, as required by the input data clusters. Moreover, when the weights are properly normalized and the output classes are normalized to 0/1, the MLP achieves the performance of the maximum a posteriori receiver, which is optimal from a classification point of view (Makhoul, 1992). In terms of mapping abilities, the MLP is believed to be capable of approximating arbitrary functions. This has been important in the study of nonlinear dynamics (Lapedes and Farber, 1987), and other function mapping problems.

MLPs are normally trained with the back-propagation algorithm (Rumelhart et al, 1986; Rumelhart and McClelland, 1987). In fact the renewed interest in ANNs was in part triggered by the existence of back-propagation. The LMS learning algorithm proposed by Widrow (Widrow and Hoff, 1960; Widrow and Lehr, 1990; Widrow and Stearns, 1985) can not be extended to hidden PEs, since we do not know the desired signal there. The back-propagation rule propagates the errors through the network and allows adaptation of the hidden PEs. Two important characteristics of the multilayer perceptron are: its nonlinear PEs which have a nonlinearity that must be smooth (the logistic function and the hyperbolic tangent are the most widely used); and their massive interconnectivity (i.e. any element of a given layer feeds all the elements of the next layer).

The multilayer perceptron is trained with error correction learning, which means that the desired response for the system must be known. In pattern recognition this is normally the case, since we have our input data labeled, i.e. we know which data belongs to which experiment.

Error correction learning works in the following way: From the system response at PE, i at iteration n , $d_i(n)$, and the desired response $y_i(n)$ for a given input pattern an instantaneous error $e_i(n)$ is defined by

$$e_i(n) = d_i(n) - y_i(n)$$

..... - 3

Where $e_i(n)$ is the error function, $d_i(n)$ is the desired output, $y_i(n)$ is the predicted output at the n^{th} iteration.

Using the theory of gradient descent learning, each weight in the network can be adapted by correcting the present value of the weight with a term that is proportional to the present input and error at the weight, i.e.

$$w_{ij}(n+1) = w_{ij}(n) + \eta \delta_i(n) x_j(n) \quad \dots - 4$$

The local error (n) can be directly computed from (n) at the output PE or can be computed as a weighted sum of errors at the internal PEs. The constant δ is called the step size. This procedure is called the back-propagation algorithm.

The total outflow for a day is taken as 100. The different type of contribution is taken and compared with the outflow. Surplus is found out by subtracting the contributions from total outflow. Figs 3 to 6 show the comparison between each type of use. The graphs are found out with the help of outflow data and the water use pattern of 1997. Model sensitivity around mean is given in Table 3 to 6.

Hydro	Industry	Domestic	Surplus
1.01	0.95	0.20	1.93
0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00
1.16	0.97	0.28	2.00
0.34	0.36	0.02	0.83
0.69	0.63	0.07	1.41
0.65	0.22	0.06	0.84
1.32	2.05	0.13	3.99

Hydro	Industry	Domestic	Surplus
1.55	0.74	0.08	2.32
0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00
1.21	0.59	0.08	1.98
0.13	0.06	0.01	0.32
0.71	0.39	0.04	1.19
1.86	0.68	0.09	1.00
3.61	1.20	0.15	3.00

Industrial	Domestic	Industry	Surplus
-------------------	-----------------	-----------------	----------------

1.55	0.01	0.06	2.07
0.00	0.00	0.00	0.00
0.00	0.00	0.00	0.00
2.92	0.03	0.27	3.28
0.12	0.00	0.00	0.01
0.14	0.00	0.00	0.22
0.46	0.00	0.01	0.71
0.46	0.00	0.01	0.38

Table – 6
Sensitivity about the mean (Konar)

Industrial	Domestic	Surplus
0.83	0.18	0.51
0.00	0.00	0.00
0.00	0.00	0.00
0.91	0.61	2.21
0.00	0.00	0.15
0.02	0.08	0.18
0.00	0.03	0.29
0.09	0.33	0.30

NOTE: Table: 3 – 6: Row – 1-8: Sensitivity about the mean with all the inputs; Level and Inflow of all the four dams respectively Maithon, Panchet ,Konar and Tillaya .

The model suggests the supply to be as per demand of the upstream reservoirs in Damodar catchment which is not applied in real scenario as depicted in Figs. 3 to 6. The model also reveals that the excess water received from Konar reservoir could be distributed to Panchet reservoir to fulfill the requirement of water for hydropower generation as shown in Fig.4 and Fig.6. The domestic water requirement for Tillaya and Konar reservoirs are higher whereas Maithon and Panchet reservoirs have maximum industrial water demand. The water requirement for irrigation which is utilized maximum in the lower Damodar catchment. The model has also prescribed the sufficient water supply for industry but priority is given to hydropower for Maithon and Panchet reservoirs. It may be noted that the water for industry is utilised to feed the small industries that are sprouting up on the adjacent lands of Konar reservoir. Tilaiya reservoir has hydropower station although small amount of electricity (4 MW) is generated but this electricity is essential for the adjacent population and some minor small scale industries. Table–7 shows the average distribution of water for a single day in summer delineated by the model (Majumder, 2006).

Table – 7
Distribution of Water for specific use of 1 single but standard day in Summer

Name of the dam	Used in Industry		Used for Domestic purpose		Used in HydroPower (at an instant)		Surplus	
	A	P	A	P	A	P	A	P
Maithon	62.33	66.85	6.92	-2.08	83.63	86.66	-52.89	-59.95
Panchet	32.91	38.27	3.66	-0.62	70.4	78.25	-6.97	-23.14
Konar	0	3.18	17.68	18.89	-	-	0	-5.68
Tilaiya	0	-0.44	0	0.07	0	-2.07	0	-3.98

NOTE: A: Actual data P: Predicted data, Bolded Text shows the type of use prioritized by the model

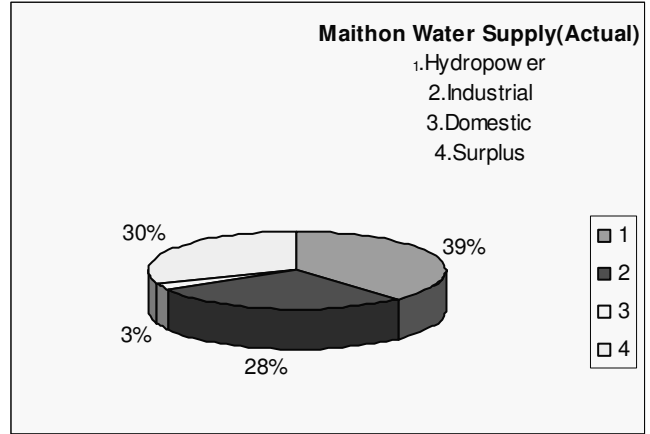
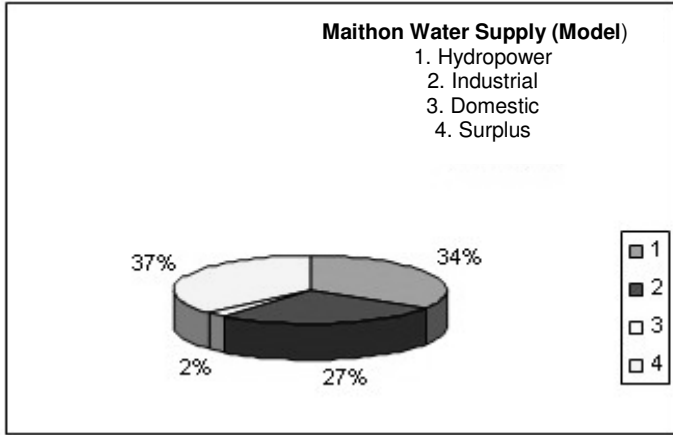


Fig. 3: Water Supply of Maithon compared with the neural model output of the same

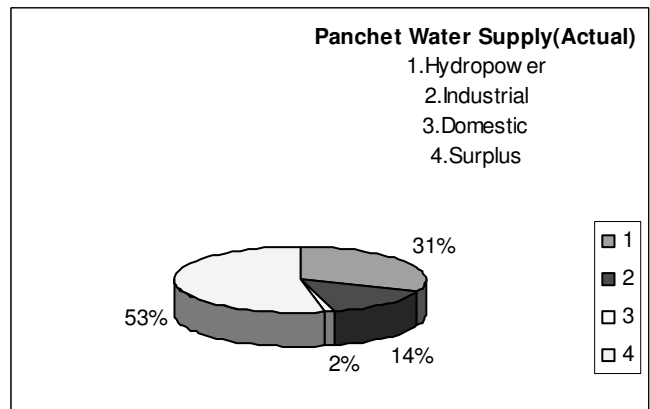
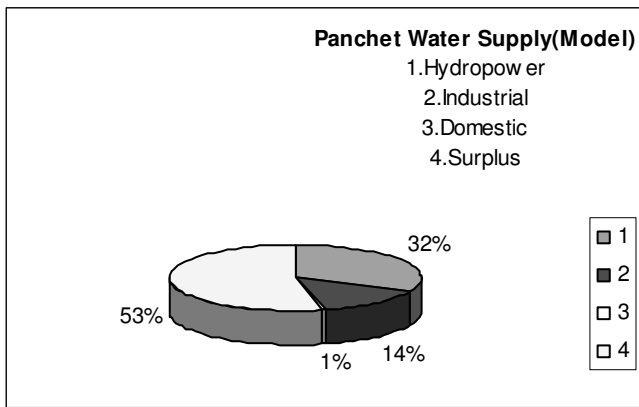


Fig. 4: Water Supply of Panchet compared with the neural model output of the same

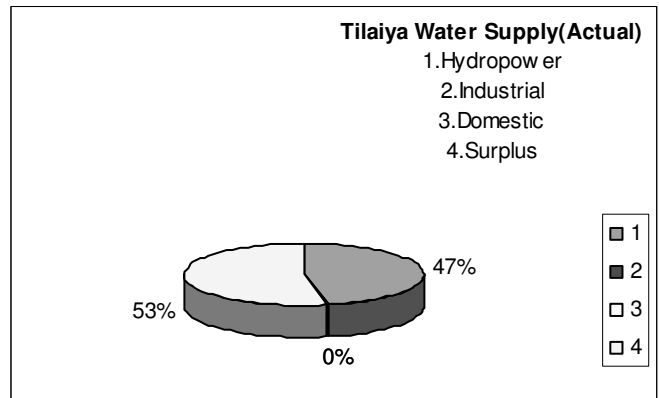
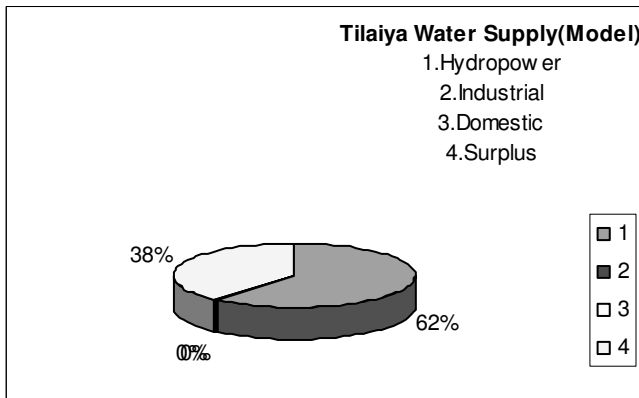


Fig. 5: Water Supply of Tilaiya compared with the neural model output of the same

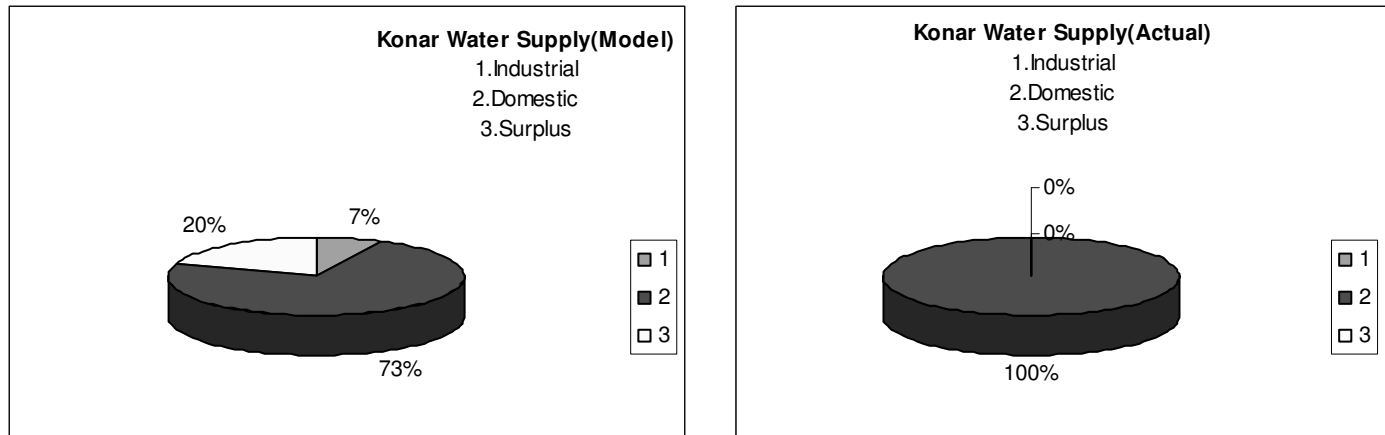


Fig. 6: Water Supply of Konar compared with the neural model output of the same

Conclusion

The developed ANN model is validated with the observed data of Damodar River and is expected to reasonably predict the fluvio-morphological variability, such as discharge, stage and sediment flow at discrete points with respects to space and time. Choosing appropriate neural network architecture and providing field data to that network for training purpose are addressed by using a constructive back-propagation algorithm with conjugate gradient descent learning. The model parameters, as well as fluvial variables, are extensively investigated in order to get accurate results. The presented ANN model is constructed by using only field river data, and it has no boundary conditions in application. The only restriction is that the model cannot estimate accurately the sediment concentration for data out of the range of the learning pattern data. Such a problem can easily be overcome by feeding the learning patterns with wide range data. The ANN can accept any number of effective variables as input parameters without omission or simplification as commonly done in the conventional models. Site engineers can calculate sediment discharge using the ANN without prior knowledge of the sediment transport theories, provided they know the bounds of the data used to generate the ANN. The adaptability to the present condition can help in proper management of the dam area. Due to paucity of time and data in the present research work it has not been possible to incorporate all the existing man-made changes such as diversions, polluted discharge etc (Sankhua, 2003) (Nash and Sutcliffe, 1970).

Reference

Lapedes A. and Farber R., (1987)

“Nonlinear signal processing using neural networks: prediction, and system modelling. ”

LA-VR-87-2662, Los Alamos,

Lippman R., (1987)

“An introduction to computing with neural nets. ”

IEEE Trans. ASSP Magazine 4.

Majumder, M. (2006)

Prédiction and Optimisation of river Damodar, an integrated multi-reservoir system by Artificial Neural Network’,

ME Thesis, School of Water Resources Engineering, Jadavpur University, Kolkata, India

Makhoul J. (1992),

- “Pattern recognition properties of neural networks.”
Proc. 1991 IEEE Workshop on Neural Networks for Signal Processing,.
- Nash, J.E., and Sutcliffe, J.V.,(1970).
“River flow forecasting through conceptual models, part-1- A discussion of principles.”,
Journal of Hydrology, 10(3), pp.282-290.
- Roy, P.K, Roy, D, Mazumdar, A,(2004)
“An impact assessment of climate change and water resources availability of Damodar river basin”,
Hydrology Journal, Indian Association of Hydrologists(IAH),Vol.27,Number 3-4, pp.53-70.
- Roy, P.K, Mazumdar, A,(2005)
“Error Function Analysis of Runoff Hydrographs of Damodar River Basin in India Simulated by Hydrologic Modelling System(HEC-HMS)”,
Bangladesh Journal of Water Resource Research,Vol.20, pp.105-117.
- Rumelhart D., Hinton G. and Williams R. (1986),
“Learning internal representations by error propagation.”
In Parallel Distributed Processing, (eds. Rumelhart and McClelland), MIT Press.
- Rumelhart D. and McClelland J.(1987) (eds.)
Parallel Distributed Processing.
Vol I, II, MIT Press.
- Widrow B. and Hoff M.,(1960)
Adaptive Switching Circuits.
IRE Wescon Rept. 4,.
- Widrow B. and Lehr M.,(1990)
“30 years of adaptive neural networks: perceptron, madaline and back propagation”
Proc. IEEE 78, 1415-1442,.
- Widrow B. and Stearns S.,(1985)
Adaptive Signal Processing.
Prentice-Hall.
- Sankhua, R N (2003)
“Neural Network in Hydrological Modelling a case study of river Brahmaputra”
IIT (Kharagpur)