www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume - 3 Issue -9 September, 2014 Page No. 8286-8296

Modelling Reservoir Operation Using Multiple Regression & Artificial Neural Network

***S.S.Khare*[1] , *Dr. A.R.Gajbhiye*[2]**
1. S.S.Khare, Assistant Engineer-II, Regional Training Centre, WRD, Nagpur, (M.S.), India.
(e mail ID - shrivan.khare@gmail.com).
2. Dr. A.R.Gajbhiye, Professor & Head, Department of Civil Engineering,
Yeshwantrao Chavan College of Engineering, Nagpur, (M.S.), India.

ABSTRACT :-

Reservoir operation frequently follows a conventional policy based on Guide curves that prescribes reservoir releases. Operating policies can be derived using system techniques such as simulation, optimisation and combination of these two. In recent years, artificial intelligence techniques like Artificial Neural Network (ANN) arisen as an alternative to overcome some of the limitations of conventional methods. In most of the studies, feed forward structure and the back propagation algorithm have been used to design and train the ANN model respectively.

A case study of Wadgaon reservoir is considered. On the basis of data available and observations simulation based Multiple regression (MLR) modelling and Artificial Neural Network (ANN) modelling is carried out . Forty seven years of 10 daily Inflow data and other relevant data is used for the analysis.

The main finding of the research is that the ANN procedure to derive the general operating policy for reservoir operation gives better and robust performance, indicating that ANN has a great potential for deriving optimal operating policy for reservoir.

Key words: - Reservoir operation, Guide curves, Wadgaon Reservoir, Simulation, Multiple Linear regression, Artificial Neural Network,.

INTRODUCTION :-

To develop and assess the application potential of the Artificial Neural Network model in attaining the reservoir operational objectives one major irrigation project “Lower Wunna project” of Nagpur district is taken as a case study. Lower Wunna project envisages two storages one across Wunna river near village Wadgaon and one across Nand river near village Sadeshwar in Umred tahasil of Nagpur district. This is a multipurpose project and is intended to cater the irrigation as well as domestic and industrial water demands of the nearby area. The Wadgaon storage has two main canals one on each bank where as the Nand storage has one main canal which serves as a feeder to Left bank canal of Wadgaon storage.

Simulation model for fortyseven years of hydrological data is developed for the Wadgaon Reservoir. Using the simulation results mathematical model using multiple linear regression (MLR) technique and Artificial Neural Network (ANN) technique are formulated and the results are also compared.

STUDY OBJECTIVE :-

The objective of the present work is to study the applicability of Artificial Neural Network (ANN) in modelling reservoir operation. This objective is attained through the following steps:

1. Hydrological data like inflow, irrigation and non-irrigation demands, and physical features of the reservoir like Area, Capacity relation with elevation, FRL. MDDL and evaporation were collected from the project authority. Reservoir simulation is carried out for forty seven years with 10 daily intervals.
2. Mathematical model using i) Multiple Linear Regression(MLR) and ii) Artificial Neural Networks (ANN) are developed using the results of

simulation analysis. Which gives the functional relationship between output (closing capacity/ closing water levels/closing area) and inputs (initial storage/level, Inflow, demands & evaporation).

3. Performance of i) multiple regression analysis and ii) Artificial Neural Networks (ANN) are compared.

APPLICATION OF ANN IN RESERVOIR OPERATION:-

Since the early 1990's there has been a rapidly growing interest among engineers & scientist to apply ANN in diverse field of water resources engineering. Raman and Chandramouli (1996) used artificial neural networks for deriving better operating policy for the Aliyer dam in Tamil Nadu. General operating policies were derived using neural network model from the DP model. The results of ANN with dynamic programming algorithm provided better performance than the other models. Jain, Das and Shrivastava(1999) used artificial neural network for reservoir inflow prediction and the operation for upper Indravati Multipurpose Project, Orissa. They developed two ANN to model the reservoir inflows and to map the operation policy. They found that ANN was suitable to predict high flows. They concluded that ANN was a powerful tool for input output mapping and can be used effectively for reservoir inflow forecasting & operation. T.R.Neelkantam et ali (2000), Chandramouli et al (2002), Cancelliere et all (2002), Oscar Dollins and Eduardo Varas (2004) , Haddad and Alimohammadi (2005), Farid Sharifi, Omid Haddad and Mahsoo Naderi (2005), Paulo Chaves and Toshiharu Kojiri (2007) , Paulo Chaves & Fi John Chang (2008), Yi min Wang at all (2009) , Amir Ali Moaven Shahid (2009,Paresh Chandra Deka and V. Chandramouli (2009), El Shafie A at all (2011), Sabah S Fayaed at all (2011), Dr.Bithin Datta (2012), T. S. Abdulkadir at all (2012) are among the others successfully studied the application of ANN in optimal operation of reservoir system. They concluded and recommended that forecasting using ANN is very versatile tool in reservoir operation.

CASE STUDY :-

In this paper Wadgaon reservoir is selected for analysis . Salient features of Wadgaon Reservoir are shown below.

Sr.No.	Particulars	Wadgaon Storage
1	Location :- Village/Tahasil/Districr	Wadgaon/Umred/Nagpur
2.	River :-	Wunna a tributary of Wardha River
3.	Catchment area :- Gross/Free	1076 Sq.Km./846 Sq.Km.
4.	Avg. Annual Rainfall	1200 mm
5.	Water availability :- i) at 75 % dependability ii) at 90 % dependability	246.394 Mm^3 (Year 1972) 141.806 Mm^3 (Year 2006)
6.	Storage capacity :- Gross / Live	152.60 / 136.00 Mm^3
7.	FRL Level / MDDL Level :-	255.100 / 248.500 m
8.	FRL Area / MDDL Area :-	16.317 Mm^2 / 7.71 Mm^2
9.	FRL Capacity / MDDL Capacity :-	152.600 Mm^3 / 36.138 Mm^3
10.	Annual Water demand :- i) Irrigation ii)Water Supply iii) Evaporation	116.086 Mm^3 31.248 Mm^3 31.281 Mm^3 } 178.615 Mm^3

RESERVOIR SIMULATION :-

Simulation is perhaps the most powerful of all the tools available to water resources system analysis. The reason of its popularity and power lies in its mathematical simplicity rather than its sophistication. It is a modelling technique that is used to approximate the behaviour of a system on the computer, representing all the characteristics of the system largely by a mathematical description.

Reservoir simulation for 10 daily intervals from year 1960 to 2006 i.e. for 47 years is carried out. The demands that can be fulfilled with 80 % success i.e. 80 out of 100 years the demand is fulfilled are arrived by trial and error by adjusting the reservoir releases. With these demands the simulation is repeated and the demands are readjusted in such a way that out of Forty seven years we get minimum 38 years as successful years. i.e. the contemplated demands will be fulfilled in at least 38 years.

The table indicates the abstract of simulation study.

Sr. No.	Year	Inflow in Mm^3.	withdrawal in Mm^3.	Spill over in Mm^3.	Deficit in Mm^3.	Remarks
1	2	3	4	5	6	7
1	1960	374.725	177.352	185.081	0.000	Success
2	1961	149.698	163.584	0.000	8.273	Success
3	1962	400.559	180.731	191.110	0.000	Success
4	1963	187.631	179.363	10.515	0.000	Success
5	1964	381.501	181.092	191.780	0.000	Success
6	1965	463.197	180.648	291.074	0.000	Success
7	1966	300.258	179.420	124.466	0.000	Success
8	1967	360.663	180.138	169.308	0.000	Success
9	1968	271.561	179.756	89.772	0.000	Success
10	1969	412.722	182.245	207.535	0.000	Success
11	1970	394.323	182.957	204.409	0.000	Success
12	1971	365.242	178.927	199.242	0.000	Success
13	1972	246.394	178.615	68.235	0.000	Success
14	1973	104.605	116.921	0.000	52.986	Failure
15	1974	192.421	176.899	0.145	0.000	Success
16	1975	196.436	178.143	20.935	0.000	Success
17	1976	525.934	180.911	334.238	0.000	Success
18	1977	263.135	181.205	81.883	0.000	Success
19	1978	128.632	156.360	0.000	19.394	Failure
20	1979	323.143	178.588	116.476	1.876	Success
21	1980	246.556	181.530	67.202	0.000	Success
22	1981	487.116	183.309	296.970	0.000	Success
23	1982	192.905	178.201	33.865	0.000	Success
Sr. No.	Year	Inflow in Mm^3.	withdrawal in Mm^3.	Spill over in Mm^3.	Deficit in Mm^3.	Remarks
24	1983	270.940	180.834	101.005	0.000	Success
25	1984	345.557	179.086	174.332	0.000	Success
26	1985	616.606	176.815	444.034	0.000	Success
27	1986	448.004	180.637	251.408	0.000	Success
28	1987	363.159	181.423	182.024	0.000	Success
29	1988	50.633	60.568	0.000	107.921	Failure
30	1989	26.664	28.089	0.000	137.340	Failure
31	1990	112.890	115.556	0.000	54.871	Failure
32	1991	268.558	180.482	63.424	0.000	Success
33	1992	352.275	178.715	187.489	0.000	Success
34	1993	55.921	69.018	0.000	97.578	Failure
35	1994	478.291	181.517	266.161	0.000	Success
36	1995	150.602	177.211	0.000	0.000	Success
37	1996	453.237	183.113	246.667	0.000	Success

38	1997	382.443	180.647	215.612	0.000	Success
39	1998	147.275	164.067	0.000	8.177	Success
40	1999	137.466	136.592	0.000	28.335	Failure
41	2000	883.197	177.267	658.386	5.406	Success
42	2001	679.032	183.801	501.049	0.000	Success
43	2002	15.352	55.084	0.000	113.785	Failure
44	2003	15.972	3.041	0.000	163.198	Failure
45	2004	185.426	165.835	1.862	8.670	Success
46	2005	340.533	180.612	141.916	0.000	Success
47	2006	141.806	173.103	0.000	2.309	Success
	Average	295.557	179.802	93.031	0.000	Success

Simulation study indicates percentage success as 81 % i.e. the simulation study is giving acceptable results. The simulation study thus forms the basis for multiple regression modelling as well as Artificial Neural Network (ANN) modelling.

MULTIPLE LINEAR REGRESSION MODEL (MLR):-

The degree of relationship existing between three or more variables is called multiple regression. Regressions models are formulated using IBMSPSS 20 software in the midst of different methods which is based on following methodology:

1. The multiple linear regression model for a response variable, y, with observed values, y1, y2, …, yn (where n is the sample size) and q explanatory variables, x1,x2,…, xq, with observed values x1i,x2i,…, xqi for i = 1, …, n is :

$$Yi=\beta 0+\beta 1x1i+\beta 2x2i +- - - - -\beta qxqi +\varepsilon i$$

2. The term εi is the residual or error for entity i and represents the foray of the observed value of the response for this entity from that expected by the model. These error terms are implicit to have a normal distribution with variance σ2.

The fortyseven years 10 daily simulation study is used to perform multiple linear regression. Initial storage/area/level , irrigation demands, non-irrigation demands and evaporation are considered as independent variables. Final area/ Final storage/ Final levels are considered as dependant variables. Multiple linear regression model for final storage (MLR Cap.), final levels (MLR Level) and final area (MLR Area) are developed.

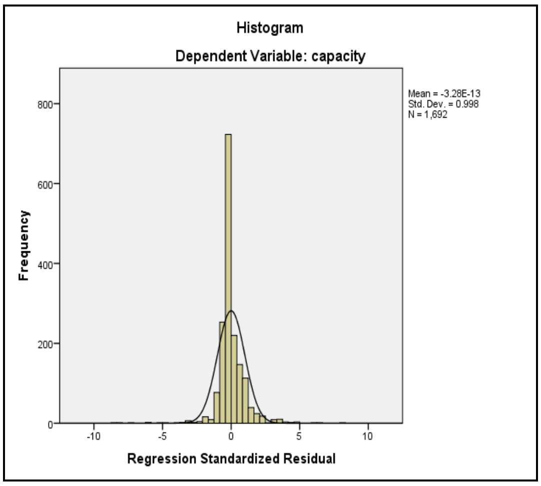
Model summary –

Model	**(R^2)**	**Adjusted (R^2)**	**F**	**Sig. F**	**Std. error of estimate**
MLR(Cap.)	0.977	0.977	10447.030	0.000[b]	7.07808
MLR (Level)	0.977	0.976	10005.822	0.000[b]	0.37845
MLR (Area)	0.981	0.981	12278.683	0.000[b]	1.43422

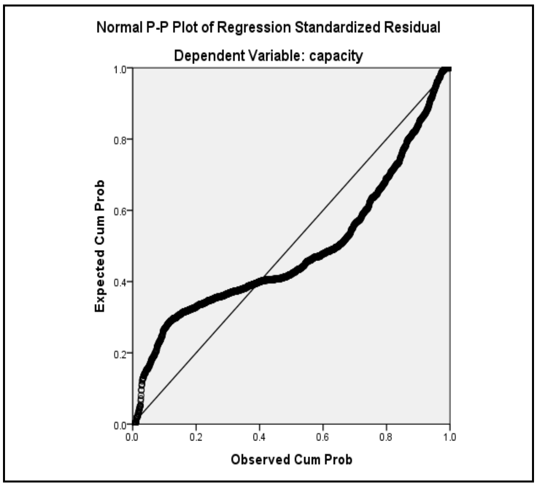
ANOVA summary –

Model		**Sum of Squares**	**df**	**Mean Square**
MLR(Cap.)	Regression Residual	3663719.727 84367.164	7 1684	523388.532 50.099
MLR (Level)	Regression Residual	10031.461 241.188	7 1684	1433.066 0.143
MLR (Area)	Regression Residual	176799.263 3463.959	7 1684	25257.038 2.057

Interpretation of MLR model for all basic variables and all readings: Table labelled ANOVA displayed the results of the analysis. The table shows the test of significance of the model using an ANOVA. There are 1691 (N - 1) total degrees of freedom. With 7 predictors, the regression effect has 7 degrees of freedom. The Regression effect is statistically significant indicating that prediction of the dependent variable is accomplished better than can be done by chance. The table labelled Model Summary provides an overview of the results. Of primary interest is the R Square and adjusted R square values, which are ranging from 0.977 and 0.981, respectively. We learn from these that the weighted combination of the predictor variables explained approximately 97 to 99 % of the variance of R.L. The prediction model is statistically significant, F (10005.822), $p < .001$, and accounted for approximately 97 to 99 % of the variance. (R^2= 0.97, Adjusted R^2= 0.98). In ANOVA table we could see the two sums of squares introduced in class – the regression and residual (or error) sums of squares. We also make assumptions about the errors. Specifically, we need to assume that the residuals are independent and normally distributed, and that they have equal variances for any predictor value. Therefore we make a normal plot . The residuals look very normal and thus the predictors mentioned in the model explained better variation in the data. Figure also depict the distribution of observed residuals matches up nicely with the distribution we would expect under normality, then residuals should fall along a straight line, as they more or less do in the plot mentioned. As deviation is substantially less from a straight line, it suggests a fewer potential deviation from normality.



Standardized Residual Plot



Normal P-P plot

ARTIFICIAL NEURAL NETWORK MODEL (ANN) :-

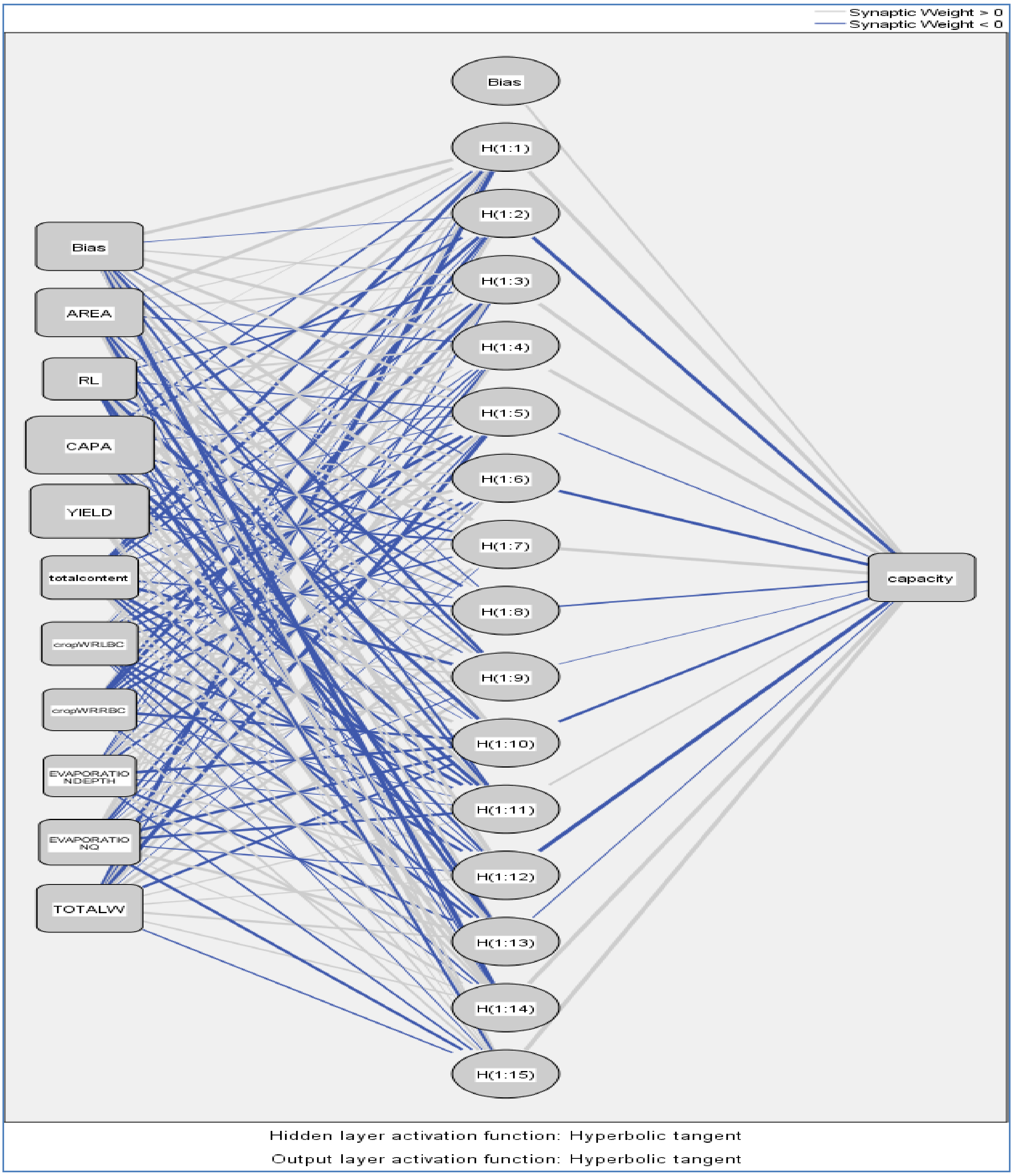
The forty seven years 10 daily simulation study is used to develop ANN Model. The standard network that is used for function fitting is a two-layer feed forward network, with a sigmoid transfer function in the hidden layer and a linear transfer function in the output layer. The number of hidden neurons is set to 15. We have tried various topologies and out of that the best results are achieved using 10-15-1 topology. The corresponding neural network model for dependent variable is best fitted with topology of 10-15-1,wherein after excluding the constant variables altogether 10 independent participating variables is considered by the network for fitting the model. With one hidden layer consisting of 15 unobservable node and one output layer is the part of topology. Training algorithm used to train the neural network is Levenberg-Marquardt (trainlm) which is recommended for most problems by various researchers. Software Matlab along with IBMSPSS 20 used for ANN modelling.

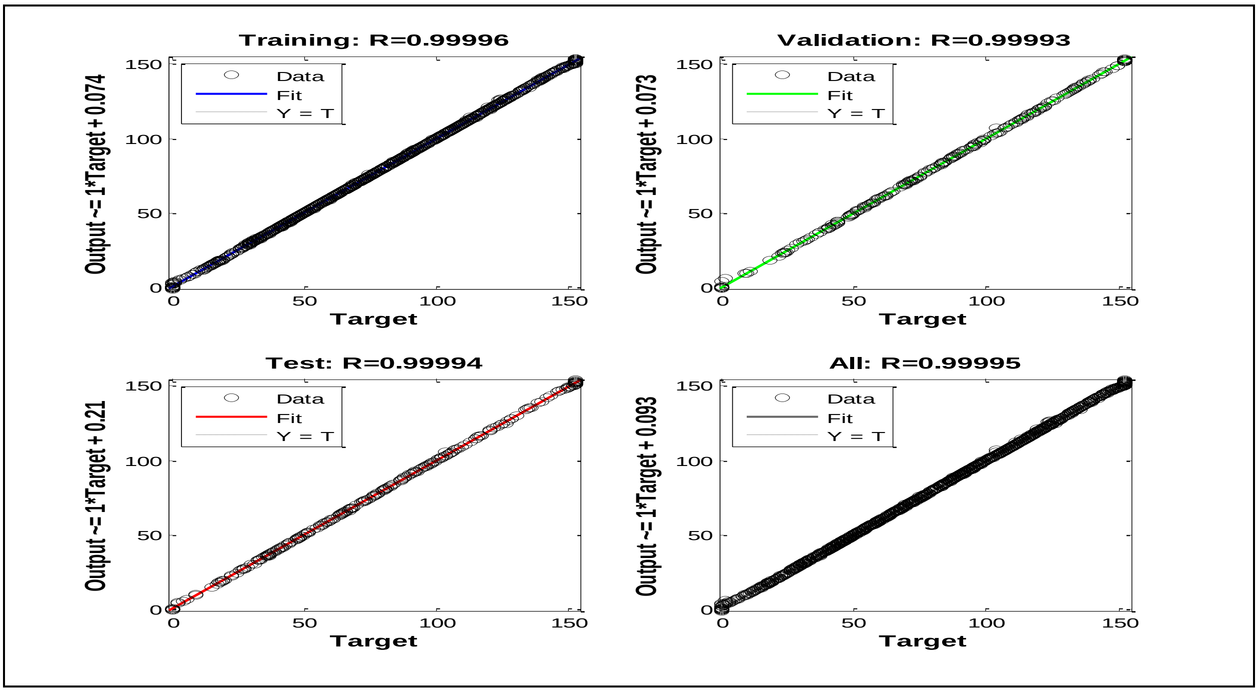
Initial storage/area/level, irrigation demands, non-irrigation demands and evaporation are considered as independent variables. Final area/ storage/level are considered as dependant variables. Multy Layer Perceptron ANN model for final storage (ANN Cap.), final area (ANN Area) and final levels (ANN Level) are developed. Hyperbolic tangent activation function is used. Out of 47 years data 33 years data is used for training and remaining data is used for testing and validation.

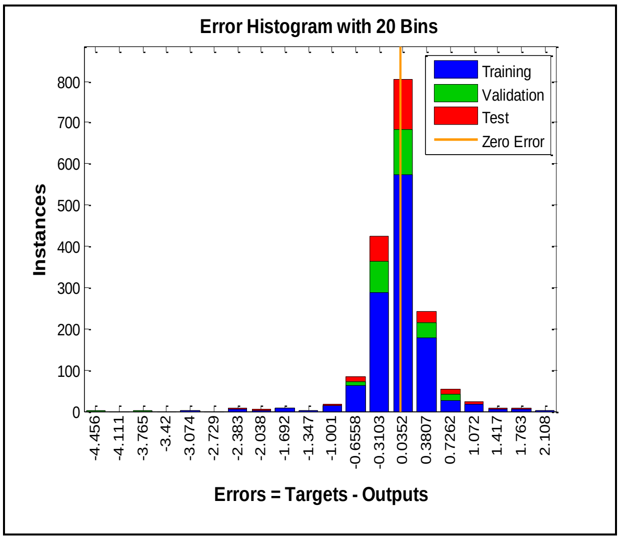
Model summary –

Model	Topology	Overall R	MSE
ANN (Cap.)	10-15-1	0.99	0.30895 at epoch 55
ANN (Level)	10-15-1	0.99	0.00851 at epoch 35
ANN (Area)	10-15-1	0.99	0.06356 at epoch 50

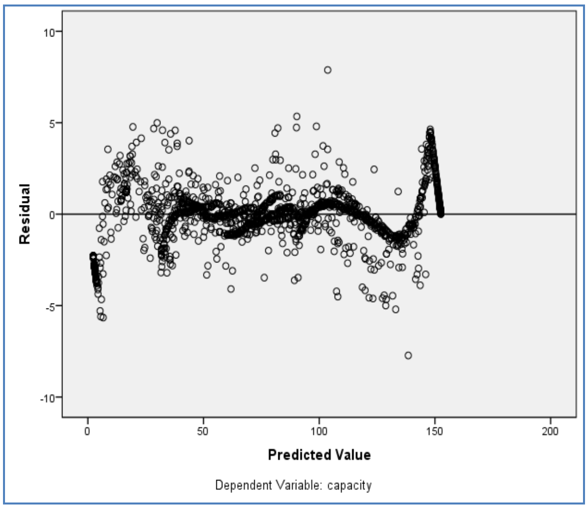
ANN network and important graphs for closing capacity (ANN Cap.) are shown below.

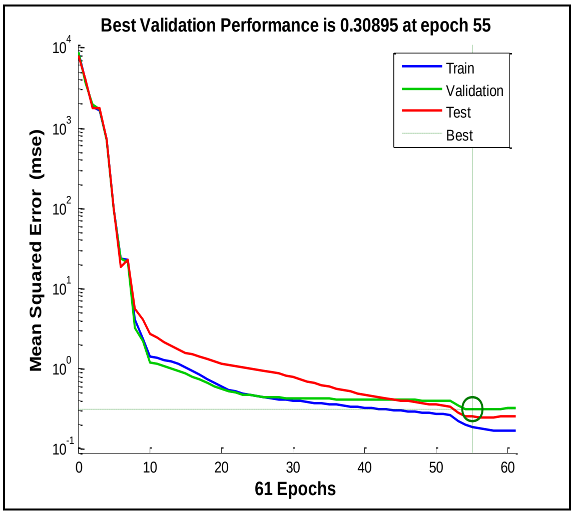


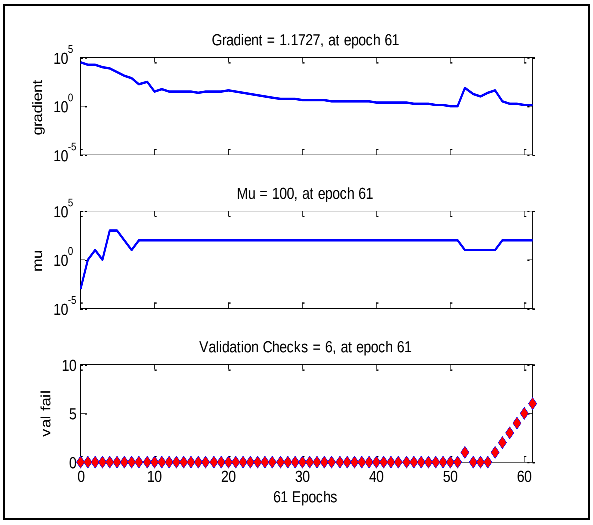




Residual Plot







Interpretation of ANN model for all basic variables and all readings: The regression plots display the network outputs with respect to targets for training, validation, and test sets. For a perfect fit, the data should fall along a 45 degree line, where the network outputs are equal to the targets. For this model, the fit is exceptionally good for all data sets, with R values in each case of 0.999 or above. If even more accurate results were required, retraining the network might be possible. Retraining will change the initial weights and biases of the network, and may produce an improved network after retraining.

The error histogram obtains additional verification of network performance. The blue bars represent training data, the green bars represent validation data, and the red bars represent testing data. The histogram can give an indication of outliers, which are data points where the fit is significantly worse than the majority of data. In this model, one can see that while most errors fall between 1.072 and -1.001, there is a training point with an error of 2.108 and validation points with errors of -4.456 and -4.111. These outliers are also visible on the testing regression plot. It is a good idea to check the outliers to determine if the data is bad, or if those data points are different than the rest of the data set. If the outliers are valid data points, but are unlike the rest of the data, then the network is extrapolating for these points. There is no significant evidence of influence of outliers on error histogram and the errors are almost normally distributed which indicates better fitting of the Neural network model.

COMPARISION OF RESULTS :-

Results of Multiple Regression model and ANN model indicates fairly equal results for all three output i.e. dependent variables namely closing area, closing levels and closing capacity at the end of each 10 days period. ANN model predicting higher values for reservoir filling period and predicting lower values for reservoir depletion period.

Reservoir Parameters for 75% dependable Year i.e. 1972(Inflow 246.394 Mm3)-

Sr.No.	Month (10 days interval)	Closing Parameters at the end of 10 days interval					
		YCapacity Regression	YCapacity ANN	Yarea Regression	Yarea ANN	Yrl Regression	Yrl ANN
1	Jun-01	31.133	30.140	14.462	13.880	249.999	249.790
2	Jun-02	28.558	26.500	12.956	12.570	249.755	249.570
3	Jun-03	27.374	25.640	12.248	12.240	249.592	249.500
4	Jul-01	39.191	51.150	14.440	17.910	250.090	251.190
5	Jul-02	65.869	78.510	19.514	22.510	251.761	252.800
6	Jul-03	92.594	108.550	25.292	29.160	253.149	254.100
7	Aug-01	124.101	145.370	31.492	34.520	254.801	254.970
8	Aug-02	151.006	151.190	35.996	35.630	255.190	255.040
9	Aug-03	151.204	150.580	35.987	35.520	255.133	255.020
10	Sep-01	150.808	151.590	35.781	35.750	255.092	255.060
11	Sep-02	149.779	151.330	35.592	35.690	255.054	255.050
12	Sep-03	144.662	148.530	34.652	35.110	254.862	255.000
13	Oct-01	140.154	143.650	33.765	34.150	254.671	254.880
14	Oct-02	133.323	137.580	32.598	33.260	254.536	254.800
15	Oct-03	126.535	129.380	31.438	32.120	254.404	254.670
16	Nov-01	120.907	122.530	30.529	31.130	254.330	254.570
17	Nov-02	115.610	115.270	29.624	29.960	254.227	254.400
18	Nov-03	110.341	108.240	28.723	28.700	254.124	254.190
19	Dec-01	107.474	104.850	28.350	27.970	254.136	254.080
20	Dec-02	104.022	100.790	27.761	27.120	254.069	253.920
21	Dec-03	100.583	96.960	27.174	26.270	254.003	253.760
22	Jan-01	95.553	92.430	26.237	25.190	253.860	253.500
23	Jan-02	88.480	86.330	24.832	24.190	253.070	252.990
24	Jan-03	83.481	81.030	23.715	23.020	252.843	252.700
25	Feb-01	79.954	77.340	22.952	22.180	252.662	252.530
26	Feb-02	76.617	73.970	22.218	21.470	252.503	252.340
27	Feb-03	73.304	70.550	21.489	20.780	252.345	252.160
28	Mar-01	68.621	67.140	20.530	19.930	252.086	252.050
29	Mar-02	65.135	63.790	19.768	19.230	251.921	251.870

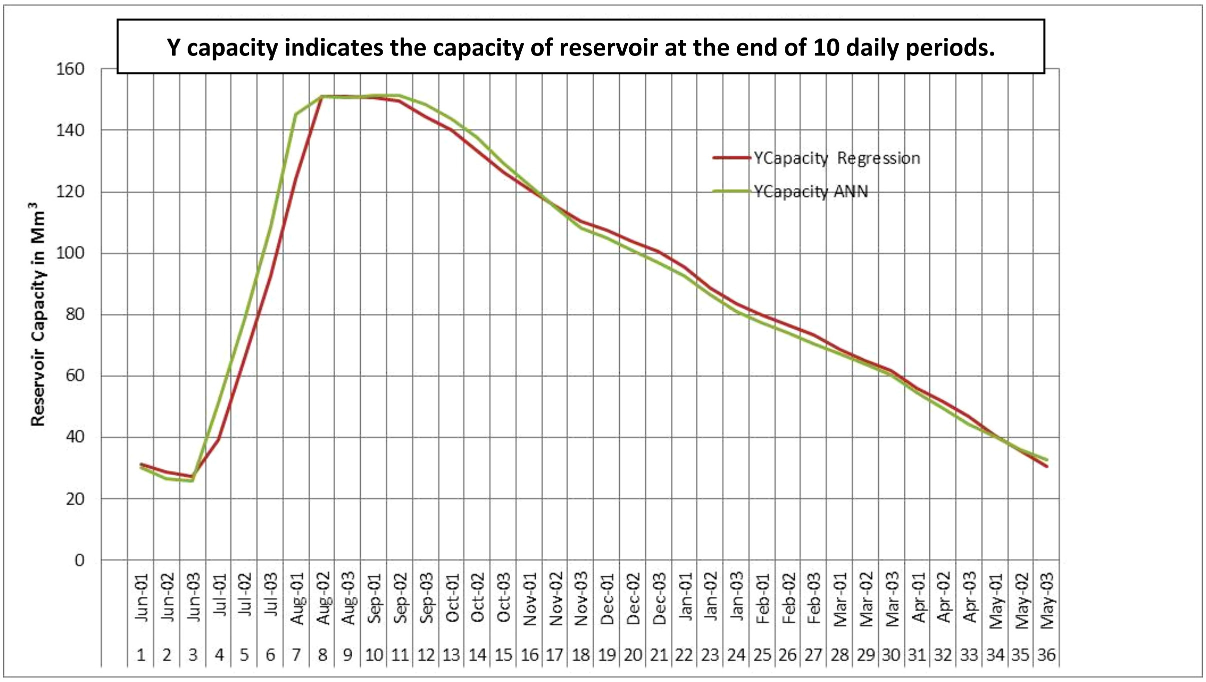
Sr.No.	Month (10 days interval)	Closing Parameters at the end of 10 days interval					
		YCapacity Regression	YCapacity ANN	Yarea Regression	Yarea ANN	Yrl Regression	Yrl ANN
30	Mar-03	61.635	60.200	19.034	18.580	251.743	251.680
31	Apr-01	56.046	54.590	17.907	17.590	251.447	251.420
32	Apr-02	51.375	49.400	16.813	16.610	251.192	251.130
33	Apr-03	46.652	44.380	15.970	15.790	250.928	250.820
34	May-01	40.536	40.190	14.954	15.150	250.562	250.510
35	May-02	35.698	35.960	14.488	14.670	250.279	250.180
36	May-03	30.627	32.750	14.190	14.380	249.992	249.900

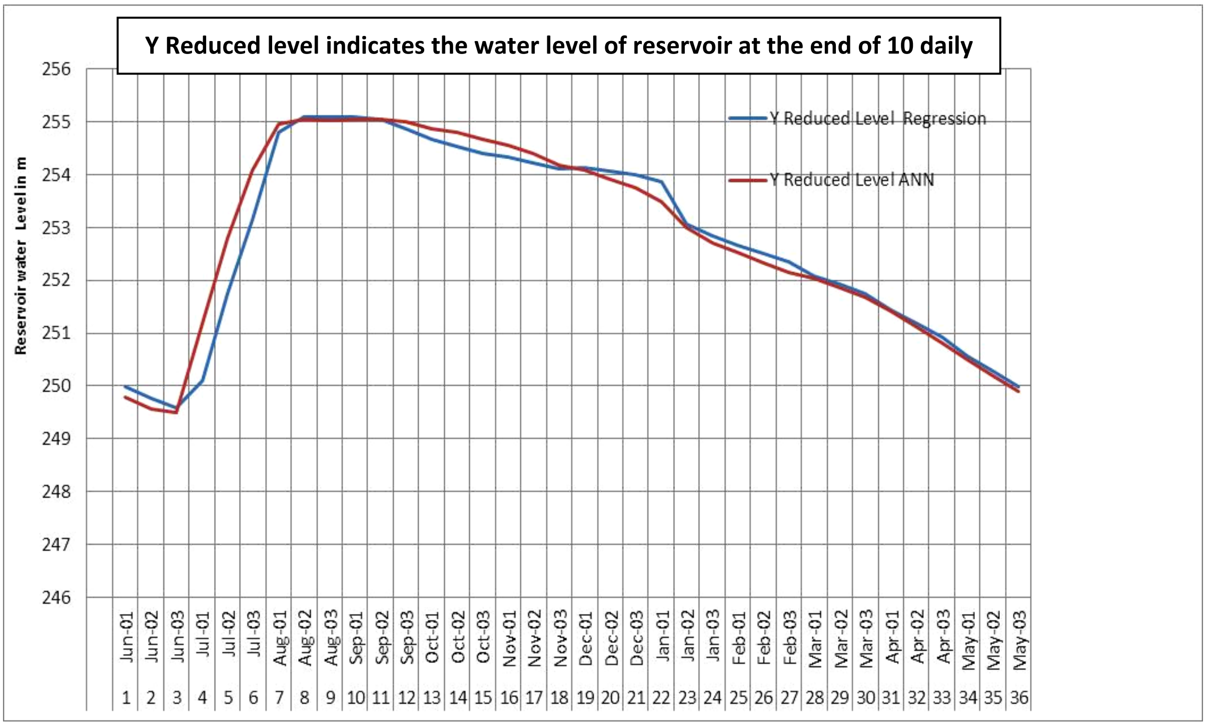
Where ;

1. Ycapacity, Yarea & Yrl indicates reservoir's closing capacity in Mm^3, closing Area in Mm^2 and closing water Level in m at the end of 10 daily periods. Further Regression and ANN indicates results for regression model and ANN models.
2. June-01 = June 1 to 10; June- 02 = June 11 to 20 & June -03 = June 21 to 30

And so on for remaining months.

Reservoir capacity as well as reservoir water levels at the end of each 10 daily period is depicted in graph below





CONCLUSION :-

The objective of this research study is to develop ANN model for operation of reservoirs and assess its application potential in attaining the objectives of reservoir operation. For Wadgaon reservoir the optimal releases for 10 daily periods are arrived at by trial and error and simulating the reservoir opening and closing conditions for 10 days intervals. Historic data of inflow for 47 years is used for simulation. Mathematical model for Multiple Linear Regression(MLR) as well as Artificial Neural Network(ANN) is developed using the forty seven years simulation study. This research study shows that the results by MLR and ANN model are fairly similar. Fine tuning and Real time forecast can be done. However, ANN predicted comparatively higher values for reservoir filling (storage built up) periods whereas the MLR predicted comparatively higher values for reservoir depletion period. However, ANN procedure to derive the general operating policy for reservoir operation gives better and robust performance. This is because the ANN approach allows more complex modelling than the MLR approach. ANN is able to produce suitable degree of nonlinearity to match the considered pattern as closely as possible, indicating that ANN has a great potential for deriving optimal operating policy for reservoir.

REFERENCES:-

1. H.Raman & V.Chandramouli (1996),” Deriving a general operating policy for reservoirs using Neural Network, Journal of Water planning and management, ASCE.
2. S.K.Jain, A. Das & S.K.Shrivastava (1999), “Application of ANN for reservoir inflow prediction & operation”, Journal of Water planning and management, ASCE.
3. T.R.Neelkantam & N.V.Pundarikantham (2000), “Neural network based simulation & optimisation model for reservoir operation”, ”, Journal of Water planning and management, ASCE.
4. V.Chandramouli and H.Raman (2001), “Multireservoir modelling with dynamic programming and neural networks”, Journal of Water planning and management, ASCE.
5. Farid Sharifi, Omid Bozorg & Mahsoo Naderi (2005), “Reservoir optimal operation using dynamic programming and Artificial Neural Network “, Proceeding of sixth WSEAS Int. Conf. on evolutionary computing, Lisbon, Portugal.
6. Paulo Chaves & Fi-John Chang (2008),"Intelligent reservoir operation system based on evolving Artificial Neural Networks", Journal of Advances in Water Resources.
7. Amir Ali Moaven Shahidi (2009),"Evaluation of combined model of DP and Neural Networks in single

reservoir operation", Journal of Applied Sciences Research.
8. Yi-min Wang, Jian-xia Chang & Qiang Huang (2010),"Simulation with RBF Neural Network model for reservoir operation rules", Journal of Water Resources management.
9. Sabah Fayaed, Ahmed El-Shafie & Othman Jaafar (2011), "Performance of ANN & regression techniques for simulation model in reservoir inter-relationships, International journal of the Physical Sciences.
10. Ozlem Terzi & Sadik Onal (2012),"Application of ANN and multiple regression to forecast monthly river flow in Turkey",African Journal of Agricultural Research.
11. S.S.Khare & A.R.Gajbhiye, (2013),"Application of ANN in operation of Reservoirs, IOSR Journal of Mechanical & Civil Engineering".

****