IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 2, February 2021

DOI 10.17148/IJARCCE.2021.10205

Prediction of Fish Production in Tamil Nadu Reservoirs Using Artificial Neural Network (ANN)

D. Karunakaran[1], M. Balakrishnan[2]

Research Scholar, Research and Development Centre, Bharathiar University, Coimbatore, Tamilnadu, India[1]
Principal Scientist, National Academy of Agricultural Research and Management, Hyderabad, Telangana, India[2]

Abstract: Fish is nutritious, cheap and healthy source of protein. Fishery sector plays a very important role for nation building by providing nutritional health security. India has very vast inland resources like rivers and canals (1,71,334 km), reservoirs (3.15 million ha), floodplain wetlands (0.24 million ha), estuaries (0.27 million ha) and ponds and tanks (2.25 million ha)[15]. Fourteen million inland farmers are getting employment and livelihood support from these resources. Among Inland resources, reservoirs are considered a main resource both in terms of surface area and offer immense scope for increasing fish production. More than 3 million ha of manmade reservoirs in the country has a potential for increase the fish production. As populations are increasing in many fold there is a greater need to obtain as much fish as possible from these resources to meet the demand. Estimation of reservoir yield is a critical component for fisheries managers to adopt suitable scientific management practices to enhance the fishery production in these resources.

Keywords: Artificial Neural Network, Inland Fisheries, Reservoir

I. Introduction

A. *Artificial Neural Networks (ANNs)*

ANNs are statistical model based on biological neural network, which has a potential to solve complex nonlinear problems. It resembles the function of brain by acquiring knowledge from the assigned interconnections (synaptic weights and biases) of processing elements (nodes or neurons) through a learning process [10]. The technique used to learn of ANNs is known as learning algorithm. It adjusts the weights associate with neurons of the network in order to predict output values based on input parameters from training values. After learning, correlate input-output data set mathematically defined as:

$$\rightarrow f{:}x_i \rightarrow z_k$$

Where x_i is a input vector and z_k is output vector indicates non-linear relation between x_i and z_k.

1) *ANN architecture types*:

The structure of ANN is dependent on optimization algorithm used to learn the network. Some of the major popular network architectures are multilayer perceptron (MLP), recurrent, radial basis function, self-organizing feature map, learning vector quantization, counter propagation, and cascade correlation etc. In general, most of the prediction or forecasting applications are being used MLP neural network architecture [2,4]. A brief description of MLP architecture is mentioned below.

A MLP is a multi-layered feed forward network consist of one input layer(i), one or more hidden layer(j) and one output layer(k). A common architecture of MLP type network is shown in Fig.1. The layers *i*, *j*, and *k* are logically

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 2, February 2021
DOI 10.17148/IJARCCE.2021.10205

interconnected with group of nodes and interlinked by adjustable synaptic weights (w_{ji}). The cycle of calculating these weights are known as 'training'. The input-hidden layer weights are represented as w_{ji} and w_{kj} denote weights of hidden-output layer.

The neuron in the hidden layer helps to recognize relevant pattern between input and output parameters. Deciding number of neurons in the network is very important criteria for getting accurate result, less neurons in the hidden layer make inadequate to learn data, where as too many nodes in the hidden layer resulted in over-fitting the training samples. The mathematical equation of neuron depicted in Fig.1 is mentioned below.

$$y_j = \sum_{i=1}^{n} w_{ji} x_i \quad (2)$$

$$v_j = y_j + b_j \quad (3)$$

Where y_j = linear combiner (summing junction); w_{ji} = synaptic weights from i to j; v_j = biased linear combiner; b_j= w_0x_0 = constant referred as threshold or bias term. The transfer function (ø) is applied to v_j, to obtain activation potential output (z_k).

$$z_k = \phi(v_j) \quad (4)$$

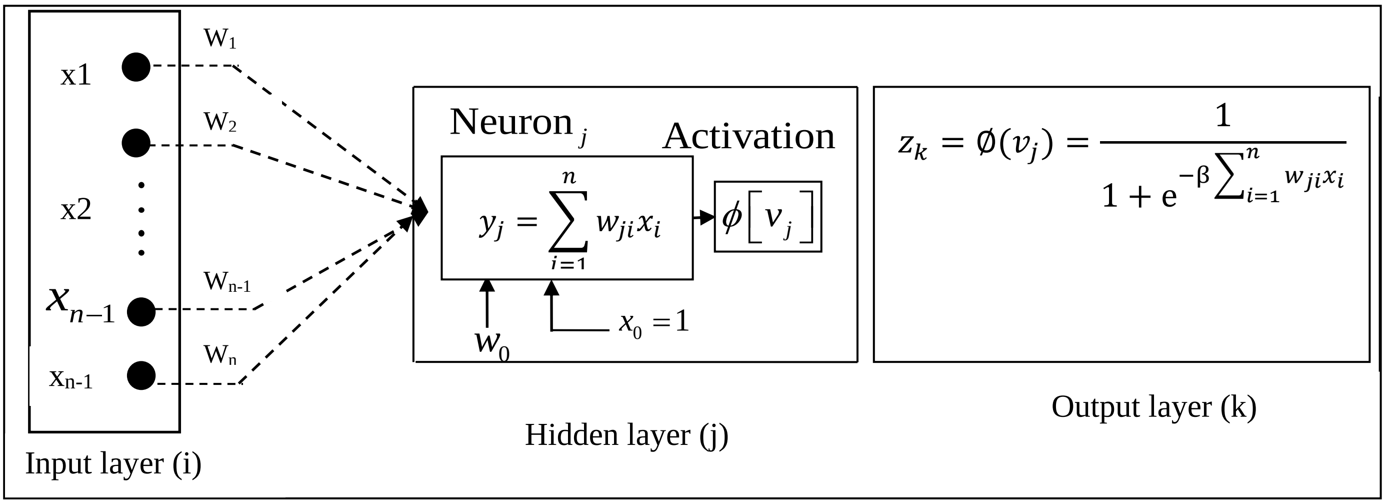

Fig. 1. A typical architecture of a MLP type neural network

2. Network training to minimize errors:

The fundamental aim of training is to improve the network's performance and minimize the system's error. There are many types of training algorithms used to train the network based on network architecture. The Back propagation (BP) algorithm is most commonly used in MLP neural networks [12]. The BP is a gradient descent optimization technique which results a suboptimal solution to the problem when it descends to a local minimum. In MLP network, the BP has mainly two passes of input propagation, i.e., forward (Eq. 5) and backward (Eq. 6).

Forward pass: $z_k = \phi\left(\text{Neuron}_j + \text{bias}\right) = \phi\left(\sum_{i=1}^{n} w_{ji} x_i + \text{b}_\text{j}\right)$ (5)

Back propagation (error) = $e_i = -\varepsilon_i + \sum_{j>1} w_{ji}\delta_j$ (6)

Where δ = summation index enforces j>1; e = vector of network produced error; ε = injected error.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 2, February 2021

DOI 10.17148/IJARCCE.2021.10205

Error (E) among target and output is computed as:

$$E = \frac{1}{2}\sum_{i=1}^{n}\left(t_i - o_i\right)^2 \quad (7)$$

Where n is the total number of training patterns, t_i is target at i^{th} node and o_i is network out at the i^{th} node. Error sends backward to the input layer from output layer for adjusting weights based on gradient equation mentioned below:

$$\Delta w_{ji} = -\eta \frac{\partial E}{\partial w_{ji}} \quad (8)$$

$$w_{ji}\left(k+1\right) = w_{ji}\left(k\right) + \Delta w_{ji}\left(k\right) \quad (9)$$

Where η is learning rate, which decides steps size at each iteration

3. Network training to update the weights:

The major challenge in training neural network is best time to stop training. Training the model with many epochs results in over fitting and at the same time too little training leads to under fitting. Generalization is the quality of neural networks that are able to provide accurate prediction with a minimum error for the unknown input variables. The main cause for the fail of generalization is due to over-fitting. To improve network training efficiency and speed the Levenberg-Marquardt(LM) is used with early stopping criteria [8, 9].

Updating of weight using LM algorithmis given by [13]:

$$W_{ji}(k+1) = W_{ji}(k) - [J^T J + \eta I]^{-1} J^T e \quad (10)$$

Where J, J^T, $J^T J$ and I are Jacobian matrix, transpose of j matrix, Hessian matrix and unit matrix respectively.

4. Multiple Linear Regression (MLR) Model:

MLR is the simple and straight forward statistical techniques; it establishes the linear relationships between two variables by fitting linear equation. It predicts dependent variable using more than one Independent variables. The principal of MLR model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots\ldots + + \beta_{n-1} x_{n-1} + \beta_n x_n \quad (11)$$

Where *y* = dependent variable, function of *n* independent variables(x);β_0 = y-intercept; β_1, ... ,β_{n-1}, β_n = partial regression coefficients of each explanatory variable.

In this study XLSTAT 2015 software tool was used to predict of fish production estimation using MLR techniques. The difference between the observed and predicted fish production gives an idea about the fit of the model and this discrepancy is called residual error.

B. Evaluation of Models' Performance

The performance of the model to predict fish production was evaluated. The indicator used for assessing models are; root mean squared error (RMSE), mean absolute error (MAE), coefficient of determination (R^2) and average target fish production values (R_{ratio}). The descriptions of above mentioned indices are mentioned below.

1. Root Mean Squared Error (RMSE):

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Vol. 10, Issue 2, February 2021
DOI 10.17148/IJARCCE.2021.10205

RMSE is used to measure difference between the observed and predicted values. Lower value of RMSE and R-Square close to 1 indicates the accuracy of model is good. It is expressed as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(T_i - O_i)^2} \tag{12}$$

Where T_i is target and O_i is output received from the models; n=number of population.

2. Mean Absolute Error (MAE):

MAE is calculated based on arithmetic average of absolute error.

$$\text{Mean Absolute Error (MAE)} = \frac{1}{n}\sum_{i=1}^{n}|T_i - O_i| \tag{13}$$

3. Coefficient of Determination (R^2):

Coefficient of Determination is indicated as R^2. It determines potential of the model to predict the dependent variable. The R^2 values are ranges from 0 to 1, higher value indicates goodness of fit of model. It is expressed as:

$$R^2 = \frac{\left[\sum_{i=1}^{n}(O_i - \overline{O})(T_i - \overline{T})\right]^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2 \sum_{i=1}^{n}(T_i - \overline{T})^2} \tag{14}$$

Where $\overline{T}$ and $\overline{O}$ are average of observed and predicted output from MLR or ANN models) respectively.

4. Ratio of Average Output to Average Target Values (R_{ratio}):

R_{ratio} is used to determine whether the models over estimated (R_{ratio}>1) or under estimate (R_{ratio}< 1).

$$R_{ratio} = \frac{\overline{O}}{\overline{T}} \tag{15}$$

C. Model Development

1. Determination of Parameters:

Learning rate is one of the very most important parameter for tuning neural network to achieve good performance. Besides learning rate other parameters used to calibrate network are activation function, error function, learning rule and initialization of weights. The training parameters used in this model were presented in Table 1. Transfer function are applied both hidden and output neurons. Most commonly used transfer function i.e 'sigmoid' and 'linear' are used in the hidden layer and output layer, respectively [14].

In this study, a three- layer network with a sigmoid transfer function in the hidden layer and linear transfer function in

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 2, February 2021

DOI 10.17148/IJARCCE.2021.10205

the output layer were used. The network training is continued till RMSE error reaching threshold. In this study initial weight range from -0.5 to 0.5 is used. These weights are then optimized for the LM training algorithm.

Table 1. Training parameters of developed models

Training parameters	Trial range	Optimum values
Learning rate	0.1 to 0.9 with a step of 0.01	0.55
Momentum	0.1 to 0.9 with a step of 0.1	0.7
Number of hidden nodes	1 to 50	20
Number of epochs	100 to 1000 a step of 50	1000
Threshold RMSE error	0.1	
Activation function	Sigmoid	

II. Materials and Methods

Secondary data of sixteen reservoirs of Tamil Nadu were used for this study. The reservoirs are classified into two groups based on geological location. Eight reservoirs are located in plain land and other group of eight reservoirs are in rain shadow region. The rain shadow reservoirs are located basically in Western Ghats (Vaigai, Manimuthar, Tirumoorthy, Palarporanthalar, Amaravathy, and Pillur) and Eastern Ghats(Gunderipallam). Thesereservoirs are getting rainfall both from south-west and north-east monsoon with less anthropogenic disturbance and get detritus load from forest litter. The other group of reservoirs (Wellington, Uppar, Orathupalayam, Krishnagiri, Vembakottai, Vidur, Odathurai) are located in plan land and elevation ranges from 38 to 2143 m above MSL and practiced intensive culture-based fisheries [16].

There are eight independent variables i.e. transparency, total phosphorus, total alkalinity, stocking (nos.), area, TDS, depth, total hardness, morpho-edapic index were used for prediction of fish production in Plain lands reservoirs. Whereas, in rain shadow reservoirs,altitude, water temperature (oC), sp-conductivity, transparency, total phosphorus, total alkalinity, stocking (nos.), area, TDS, depth, total hardness, morpho-edapic index are explanatory variables used to predict fish production.

III. Results and Discussion

A. Neural Network prediction model for Plain reservoirs in Tamil Nadu

The ANN model performed well with low RMSE and MAE and high R^2 values as compared to MLR models. Performance indices of Artificial Neural Network and MLR models for prediction of fish production of Plain reservoir are mentioned in table 2.

Table 2. Performance indices of Artificial Neural Network and MLR models for prediction of fish production (Plain)

Performance Indices	ANN	MLR
RMSE	8.470	21.164
R^2	0.991	0.945
R_{ratio}	0.977	1.000
MAE	4.561	12.861

The scatter plots of ANN and MLR predicated fish production with respect to observed fish production are mentioned in Fig. 2 and 3 respectively. ANN predicted fish productions better than MLR with the observed fish production. While comparing Fig 2, Fig 3 indicate that the line of best fit is better in ANN compared with MLR. The estimation of fish production byANN were very accurate with high values of R^2 (= 0.991).

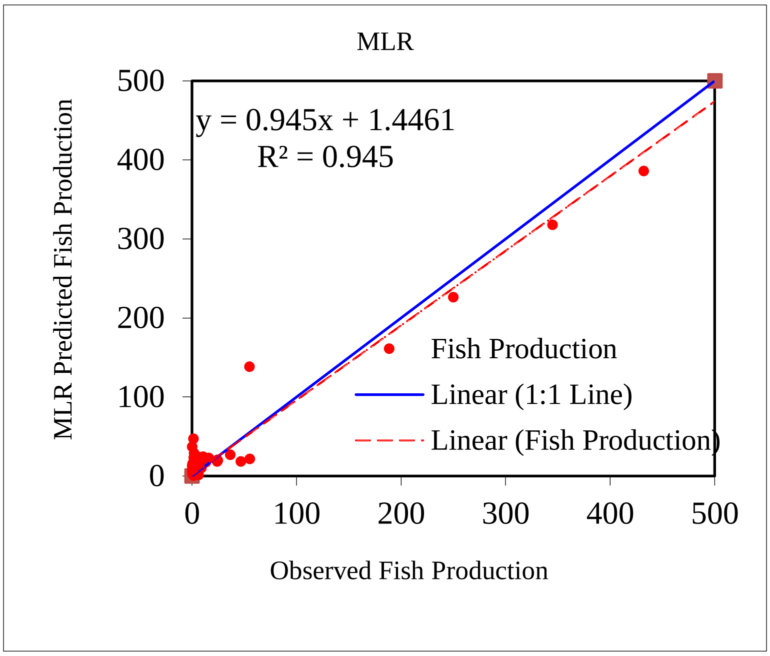

Fig. 2. Scatter plot of MLR vs.observedfish production

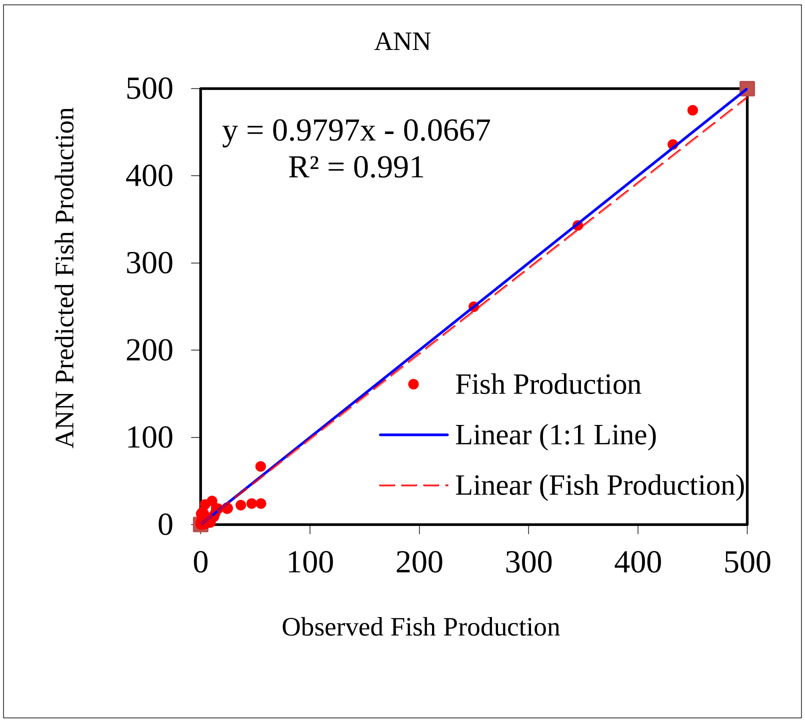

Fig. 3. Scatter plot of ANN vs observed fish production (Plain)

B. Neural Network prediction model for Rain Shadow Reservoirs

The estimates of RMSE, MAE, R^2 and R_{ratio} for ANN model were good with values 19.378, 7.658, 0.907, and 1.001 as compared to MLR model. The ANN model performed well with low RMSE and MAE and high R^2 values as compared to MLR models.

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 2, February 2021

DOI 10.17148/IJARCCE.2021.10205

Table 3. Performance indices of ANN and MLR Model

Performance Indices	ANN	MLR
RMSE	19.378	37.824
R^2	0.907	0.621
R_{ratio}	1.001	1.000
MAE	7.658	20.954

Fig 4 and 5 shows that scatter plots of estimated fish production with respect to observed fish production generated from ANN and MLR respectively. While comparing Fig 4, Fig 5 indicate that the estimation of fish production by ANN were very much accurate with high values of R^2 (= 0.907). From the above analysis results shows that estimation of fish production by the ANN gave better estimation compared to MLR models.

.

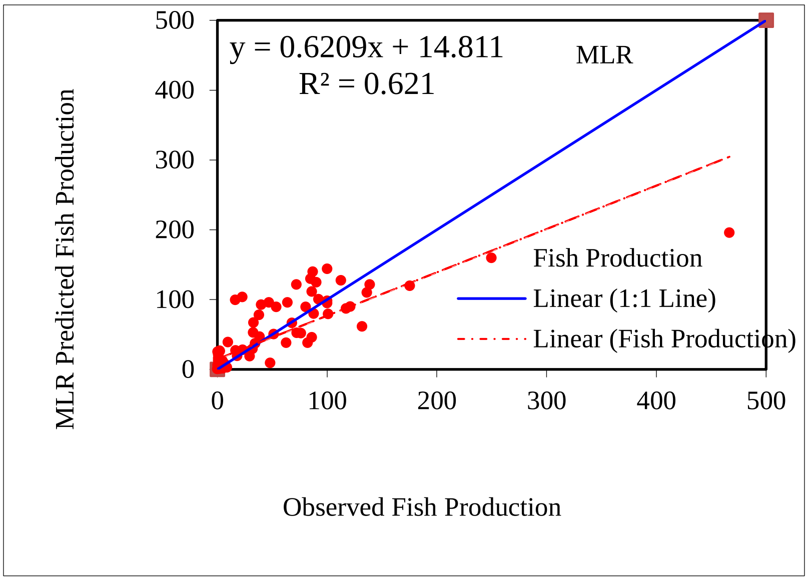

Fig. 4. Scatter plot of MLR vs observed fish production (Rain)

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 2, February 2021

DOI 10.17148/IJARCCE.2021.10205

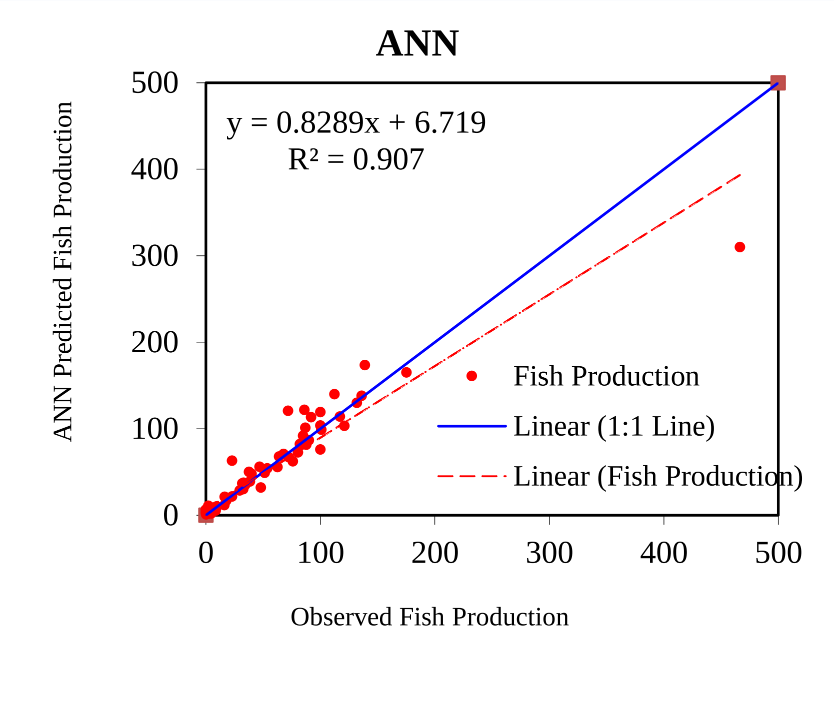

Fig. 5. Scatter plot of ANN vs observed fish production (Rain)

IV. CONCLUSION

This study demonstrated prediction of fish production using backpropagation algorithm using Neural network with environmental parameters. The independent variables for reservoirs located in plan area are different from reservoirs located in rain shadow reservoirs, where culture-based fisheries are in practice. India comprises wide range of ecological and weather conditions. Prediction of fish production potential in reservoir is an important for fisheries managers to take up suitable management practice as per regional environment condition to enhance the fish production. Artificial Neural Network gave better estimation compared to MLR models and helps to understand behaviour of the reservoir to formulate appropriate regional specific management practices to improve fish production in reservoir.

REFERENCES

[1]. D. Karunakaran, M. Balakrishnan (2019). Prediction of Fish Yields in Lakes and Reservoirs from simple Empirical Models using Artificial Neural Network (ANN) : An Review *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*. January-February-2019 ; 5(1) : 88-100.

[2]. Adamala, S.,Raghuwanshi, N. S., Mishra, A., Tiwari, M. K. (2014a). Evapotranspiration modeling using second-order neural networks. *Journal of Hydrologic Engineering*, 19(6):1131-1140.

[3]. Balakrishnan, M., Meena, K., Sethi, S. N., & Sarangi, A. N. (2007). Neural network and its application in aquaculture. *Bioinformatics and statistics in Fisheries Research*, *3*, 145-151.

[4]. Adamala, S.,Raghuwanshi, N. S., Mishra, A., Tiwari, M. K. (2014b). Development of generalized higher-order synaptic neural-based ET_o models for different agro ecological regions in India. *Journal of Irrigation and Drainage Engineering*, 140(12):DOI: 10.1061/(ASCE)IR.1943-4774.0000784.

[5]. Kumar, M. N., & Balakrishnan, M. (2016). Prediction of crop yield using weather and climate parameters for sugar cane yield in India. Indian Streams Research Journal.

[6]. Balakrishnan, M., & Parthiban, C. (2013). Development of an Expert System for Dimensional and the Resolution of Soil Texture using Data Mining Concept. *International Journal of Advanced Research in Computer Science*, *4*(9), p56-59.

[7]. Kumar, M. N., & Balakrishnan, M. (2018). Sugar Cane Crop Yield Estimation Using K-Nearest Neighbors.*JRAR-International Journal of Research and Analytical Reviews (IJRAR)*,6, 1141-1148

ISSN (Online) 2278-1021
ISSN (Print) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Vol. 10, Issue 2, February 2021

DOI 10.17148/IJARCCE.2021.10205

[8]. Admala, S., Raghuwanshi, N. S., Mishra, A. (2015a). Generalized quadratic synaptic neural networks for ET_o modeling.*Environmental Processes,* 2(2):309-329.

[9]. Adamala, S., Raghuwanshi, N. S., Mishra, A., Tiwari, M. K. (2015b). Closure to evapotranspiration modeling using second-order neural networks. *Journal of Hydrologic Engineering*,20(9):07015015.

[10]. Adamala, S. (2015). Evapotranspiration and evaporation modeling using higher-order neural networks. PhD Dissertation, Indian Institute of Technology, Kharagpur, West Bengal, India.

[11]. Haykin, S. (1998). Neural Networks-A Comprehensive Foundation, 2nd ed. Prentice-Hall, Inc., Division of Simon and Schuster One Lake Street Upper Saddle River, NJ, United States

[12]. Rumelhart, D. E., McClelland, J. L. (1986). Parallel distributedprocessing: Explorations in the microstructure of cognition. MIT Press,Cambridge, MA.

[13]. Gupta, M. M., Jin, L., Homma, N. (2003). Static and dynamic neural networks: From fundamentals to advanced theory. Wiley/IEEE Press,NY.

[14]. Maier, H. R., Dandy, G. C. (1996). The use of artificial neural networks for the prediction of water quality parameters. *Water Resources Research*, 32(4):1013-1022.

[15]. Sinha, M. 1999. Vision of inland fisheries of India of twenty first century. p. 154-168. In: S.H. Abidi, N.K. Thakur, R.S. Birader and L. Shenoy.(eds.) *Vision on Indian Fisheries of 21st Century*. Central Institute of Fisheries Education, Bombay.

[16]. Murugesan. V.K, Rani Palaniswamy and Monoharan, S. Ecology and Fisheries of Selected Reservoirs in Tamil Nadu, Bul. No. 117, *Central Inalnd Fisheries Research Institute,* Bul. No. 117, 2003