International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 13, Issue 9, 2020

# Predicting the Monthly Demand for Blood Using Artificial Neural Networks: A Case Study of a Blood Bank in Nasiriyah City

**Ahmed Karim Jasim Jasim[a], Kadhim Ahmed Jawad Abbas[b],** [a]College of Administration and Economics, University of Thi-Qar,Iraq, [b]College of Administration and Economics, Mustansiriyah University,Iraq, Email: Ahmed.kareem@utq.edu.iq, uadh_jsheme@uomustansiriyah.edu.iq

The study is based on the application of the Artificial Neural Networks (ANNs) model to predict blood demand in the central blood bank at Al-Haboubi Hospital in Nasiriyah City of southern Iraq. The model network was installed, validated and tested using blood incoming and outgoing data monthly from January 2013 to December 2017. Through trial and error procedures, the number of hidden neurons was estimated. The mean square error (RMSE), mean absolute error (MAE), and correlation coefficient (R), was applied to select the best model. Artificial neural networks are considered an effective technique to estimate incoming and outgoing blood, which is very helpful for those who work in this important field.

**Key words:** *Forecasting, Artificial Neural Networks (ANNs), Blood bank.*

## Introduction

Blood is the human elixir of human life. It is responsible for many vital processes in the human body, including respiration, transportation, excretion, regulation, and defence. The blood volume in the human body ranges from 4.7 to 5 litres, which varies from person to person. In special cases, a person needs to compensate, increase or reduce the amount of blood due to illness, injury or surgery, etc. This procedure is usually performed in hospitals, clinics and specialised health centres because blood is a vital substance consisting of a number of ingredients. In special and common cases, the demand for blood is urgent and depends on the life of the injured or the patient. For the purpose of rapid response to such cases, the blood is usually collected in establishments specialised in its preservation like blood banks, which are usually part of or close to hospitals and health centres. Blood banks usually follow advanced systems and methods in the circulation and preservation of blood stocks to ensure the best use of this biological material, and rely on several methods to predict the supply and demand of blood.

Forecasting is one of the most widely used concepts in many contemporary sciences and applications, from forecasting weather conditions to predicting coffee prices to predicting demand for blood bags. Forecasting is defined as "predicting or anticipating what will happen in the future" (Russell &Taylor, 2011 : 496 ). It is defined by Schroeder & Goldstein (2018: 86) as "the art and science of anticipating future events." A forecast is also known as "an estimate of future demand, based on past demand" (Kumar & Suresh, 2008: 110). The forecasting process permeates all functional areas of the organisation. Functional unit officials rely on expectations for the formulation and implementation of their plans. Predicting is important for business plans, annual plans, and budgets. Human resource management uses forecasting to anticipate or predict recruitment and training needs. Operations managers and supply chain managers also use forecasting in the planning of production levels, the procurement of materials, labour force, production schedules, inventory, capacity, etc. Marketing, finance and other functions also involve predicting, when managers make forecasts on many variables that affect future demand such as competitors' strategies, organisational changes, technological changes, processing times, quality losses, etc. (Krajewski et al., 2013: 465). There are many ways of forecasting and the organisations often use qualitative forecasting techniques based on past opinions and experience and experience to come up with the best possible guesses. There are also a few quantitative forecasting methods available to assist managers in making planning decisions, assessing trends and predicting the future. It is worth saying that a technique that does not lead to accurate forecasting or prediction can be utilised in providing reliable guidelines or principles in decision making (Slack et al, 2010: 170). There are six basic steps in the prediction process (Stevenson, 2018: 79):

1. Determine the purpose of the forecasting process and how it is used and what is being used. This step provides an indication of the level of detail required in the forecasting, the amount of resources (personnel, equipment, money) to be used, and the level of accuracy required.
2. Determine a time horizon: forecasts should indicate a period, considering that the accuracy decreases as the time horizon increases.
3. Obtain, filter and analyse appropriate data: access to the necessary data involves considerable effort once obtained. In addition, it needs to be filtered to eliminate extreme values and incorrect data before analysis.
4. Determine the method of prediction.
5. Creation of forecasts.

6. Monitoring of forecast errors: expected errors should be monitored to determine whether the projections are operating satisfactorily, if not, the method and assumptions should be re-examined, and data validated.

**Artificial Neural Networks (ANNs)**

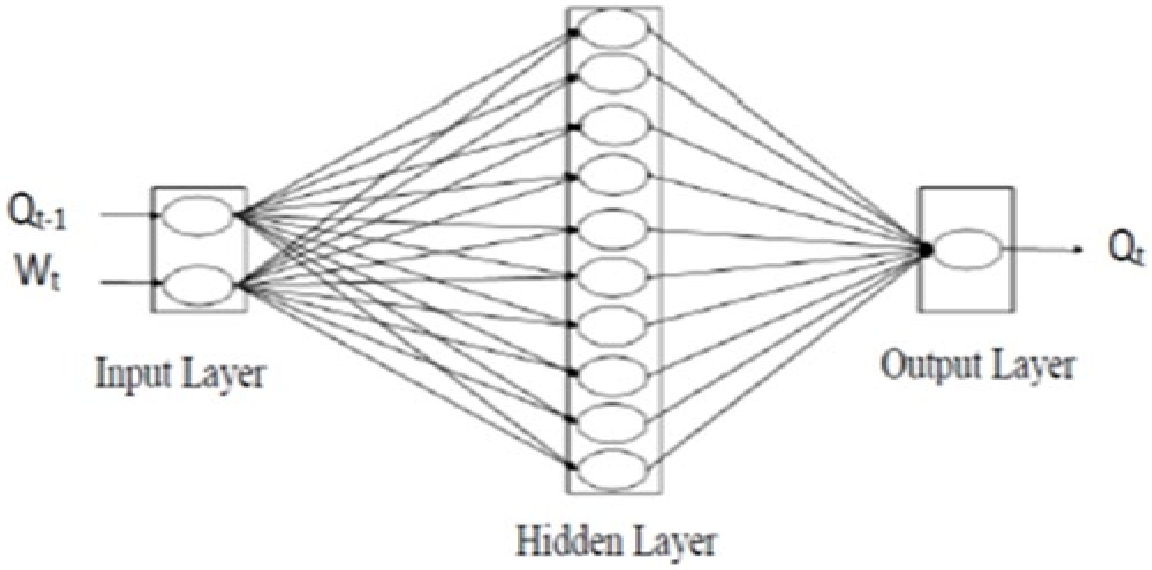
The artificial neural networks are a "collection of artificial neurons which are inspired by biological systems of living creatures". ANNs have specific characteristics that are somewhat similar to the neuronal biological networks of the human mind, and it is considered as a distributor and treatment of information in a wide range (Masrur Ahmed,2017).

ANNs refers to a widely distributed processor that has a natural propensity to store experiential knowledge and provide it to others. It is characterised by its similarity with the human mind in two ways: the process of storing knowledge, and how it is through learning, that knowledge is acquired through the network. This is done through the use of points of contact between neurons called inter locking weights. (Masrur Ahmed,2017).

Studies indicate that ANNs have been applied in various fields such as mathematics, medicine, engineering, economics and neuroscience, as it can be used in voice and speech recognition, as well as in military fields when defining targets and detecting explosives in passenger bags, and in the field of prediction. It was used in metal detection and convection prediction, and was also used in the field of operation, as it can be a model for prediction through multidimensional data collected by sensors.

The model of a network is showed in figure (1). The most efficient neural network training algorithm is known as a back propagation algorithm (BP) (Haykin, 1999). The back propagation algorithm in this paper is used to calculate nonlinear multilayer networks. The ANNs parameters (weights and biases) can be adjusted to minimise the sum of the squares of the differences between the actual values and network output values (Rankovic et al.,2010)

**Figure 1.** Architecture of multilayer perception of Artificial Neural Networks (Khairi,2017)




In any ANNs The nerve cells are called the elements through which data are processed, as these cells are grouped into layers. There are layers of neurons from the first level, which are called the input layer, and are tasked with receiving the input vector and then transfering the values to the neurons through links. This process continues until the output layer is reached. There are two types of ANNs depending on the number of layers. There are single bilayer and multi-layer networks, which can be classified in the front and back feeding networks based on the information and processing path (Haddad et al.,2005). In order to reduce the square heterogeneity between the required outputs and the response of the treatment elements, the connection weights are adjusted here. The inverse of the input-automatic correlation matrix (R-1) yields the ideal weights as well as the cross-correlation vector (P) between the inputs and the required response.

Equivalent research technology is used to obtain the minimum square performance (wi), which is equivalent to the analytical solution to this problem. This is done by applying the gradient descent by way of adjusting weights each period (Haykin, 1999):

$w_i(k+1)=w_i(k)-\eta\nabla Ji(k) \qquad \nabla Ji=\partial J\partial wi(1)$

where:

$\eta$: coefficient of learning rate.

$\nabla$(k): gradient vector of the performance surface at iteration (k) for the $i^{th}$ input node.

Performance surface ($J$) is calculated by equation (2):

$J=\Sigma_p(d_p-y_p)^2 and \quad \min J\rightarrow w_{opt}=R^{-1}P$ (2)
where:

$w_{op}$: Optimal weight ,$_p$: Target output, $y_p$: Calculated output of the p$^{th}$ output neuron.

The three layer feed forward neural network used in this study has been widely used for modelling, because these layers are sufficient to generate arbitrarily complex output signals (Lippmann,1987). The following equation is used to extract the value of multi-layer visualisation output (Nourani and Babakhani,2013):

$$y_k = f_{\circ}[\sum_{i=1}^{M_N} W_{Kj}\,.\,f_h\big(\Sigma_{i=1}^{N_N} W_{ji}X_i + W_{j\circ}\big) + W_{K\circ}](3)$$
where:

$W_{ji}$: A weight in the hidden layer connecting the i$^{th}$ neuron in the input layer and the j$^{th}$ neuron in the hidden layer.

$W_{jo}$: The bias for the j$^{th}$ hidden neuron.
$f_h$: The activation functions of the hidden neuron.
$W_{kj}$: A weight in the output layer connecting the j$^{th}$ neuron in the hidden layer and the k$^{th}$ neuron in the output layer.
$W_{ko}$: The bias for the k$^{th}$ output neuron.
$f_o$: The activation functions for the output neuron.
$X_i$: i$^{th}$ input variable for input layer.
$y_k$: computed output variable.
$M_N and N_N$: The number of the neurons in the input and hidden layers, respectively.

**Study Area and Data Set**

Nasiriyah represents the centre of the province of Dhi Qar, which is located in the south-east of Iraq on the Euphrates River. It is the fourth city in populated after Baghdad, Basra and Mosul.

Initial data for this study was collected from the expense statistics of the blood bank in Nasiriyah city. Monthly data was collected during a period of five years (January 2013 to December 2017). Table (1) shows summary statistics of the raw data.

**Table 1:** Summary statistics of the raw data for the blood bank.

| | Standard Deviation | Skewness Coefficient | Excess Kurtosis | Median (litre) | Minimum (litre) | Maximum (litre) | Average (litre) |
|---|---|---|---|---|---|---|---|
| Expense of blood | 361.6013 | -0.30321 | -0.2396 | 2261.5 | 1189 | 2996 | 2221.95 |
| Donor of blood | 300.8423 | -0.7344 | 0.001639 | 2090 | 1163 | 2562 | 2026.017 |

**Methodology**

In the study training of ANN models of different architectures, an automatic performance analysis of the networks was applied based on the correlation coefficient (R), mean squared error (MSE) and root mean squared error (RMSE) performed (Ali, 2017: 2647)

$$\mathrm{R} = \frac{\sum_{j=1}^{n}[(Y_j-\bar{Y})(\hat{Y}_j-\bar{\hat{Y}})]}{\left[\sum_{j=1}^{n}(Y_j-\bar{Y})^2 \sum_{j=1}^{n}(\hat{Y}_j-\bar{\hat{Y}})^2\right]^{1/2}} \& R = \sqrt{R^2} \quad (1)$$

$$\mathrm{MAE} = \frac{\sum_{j=1}^{n}|Y_j-\hat{Y}_j|}{n} \quad (2)$$

$$\mathrm{RMSE} = \left(\frac{\sum_{j=1}^{n}(Y_j-\hat{Y}_j)}{n}\right)^{1/2} \quad (3)$$

where:

$Y \& \hat{Y}$: The observed and estimated values respectively.
n: The number of observations.
$\bar{Y} \& \bar{\hat{Y}}$: The mean of observed and estimated values.

Three layers feed forward networks with sigmoid hidden neurons and linear output neurons are used in this study. Several networks with different numbers of hidden layer nodes (1-20) and with different transfer functions were developed. The network is trained with Leven berg-Marquradt back-propagation algorithm. The data set is scaled by using map min max function according to this scale. The ranges of the input lies inside the range ($-1 \leq x \leq 1$). Hence the total number of observations is120 samples for each model. These observations are divided into three statistical parts. 70% (84 samples) is for training. These are presented to the network during training, and the network is adjusted according to its error. 15% (18 samples) is used in validation. These are used to measure network generalisation, and to halt training when generalisation stops improving. The last part of the data set is testing at 15% (18 samples).

International Journal of Innovation, Creativity and Change. www.ijicc.net
Volume 13, Issue 9, 2020

These have no effect on training and so provide an independent measure of network performance during and after training. The early stopping method is selected to overcome the over fitting problem. A trial and error procedure based on root mean square error (Eq.3), mean absolute error (Eq.2) and coefficient of correlation (Eq.1) are used to select the best network architecture and for predicting the expenses of the blood bank.

## Results and Discussion

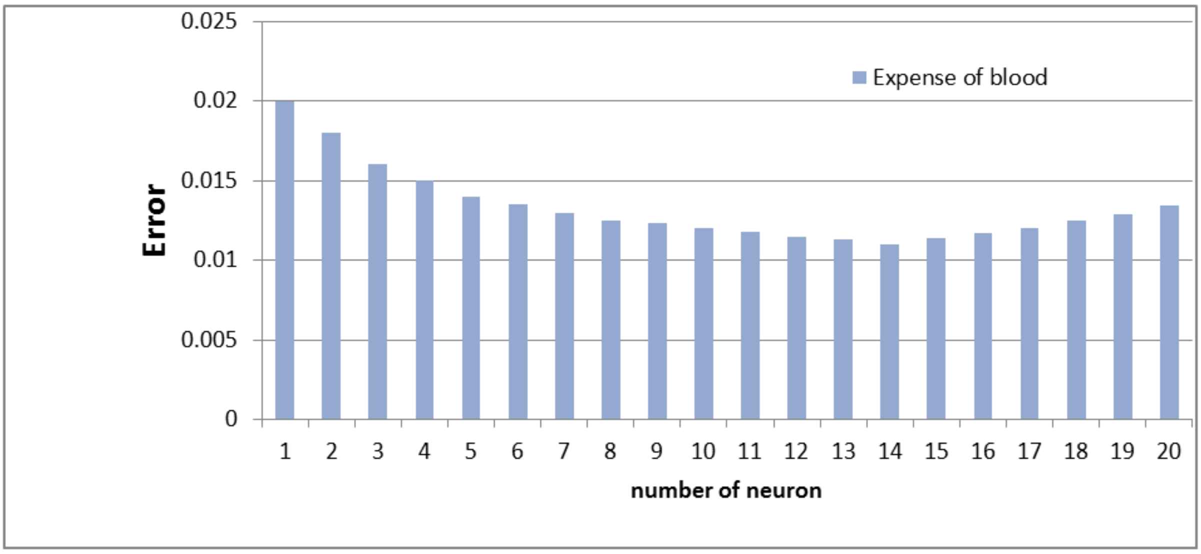
To predict the expense of blood value is used before donor of blood value. There are four models adopted for selecting the best one. Each model is described as follow:

M1: $EX_t = f(D_t)$ (7)
M2: $EX_t = f(D_t, EX_{t-1})$ (8)
M3: $EX_t = f(D_t, D_{t-1}, EX_{t-1})$ (9)
M4: $EX_t = f(D_t, D_{t-1}, D_{t-2}, EX_{t-1}, EX_{t-2})$ (10)
where:

$EX_t$: Expense of blood at a specified time.
$EX_{t-1}$ &$EX_{t-2}$: Expense of blood at t-1 and t-2 respectively.
$D_t$: Donor of blood at a specified time.
$D_{t-1}$&$D_{t-2}$: Donor of blood at t-1 and t-2 respectively.

Root mean square error (RMSE), mean absolute error (MAE) and coefficient of correlation (R) are used to evaluate the performance of models. The ANNs models were trained utilising different numbers of neuron in the hidden layer, for training, testing and validation. The results showed the error values obtained for the expense of blood when compared with the observed data shown in figure (2), for the validation data set.

**Figure 2**. Variation of error values with hidden layer neurons for the validation data



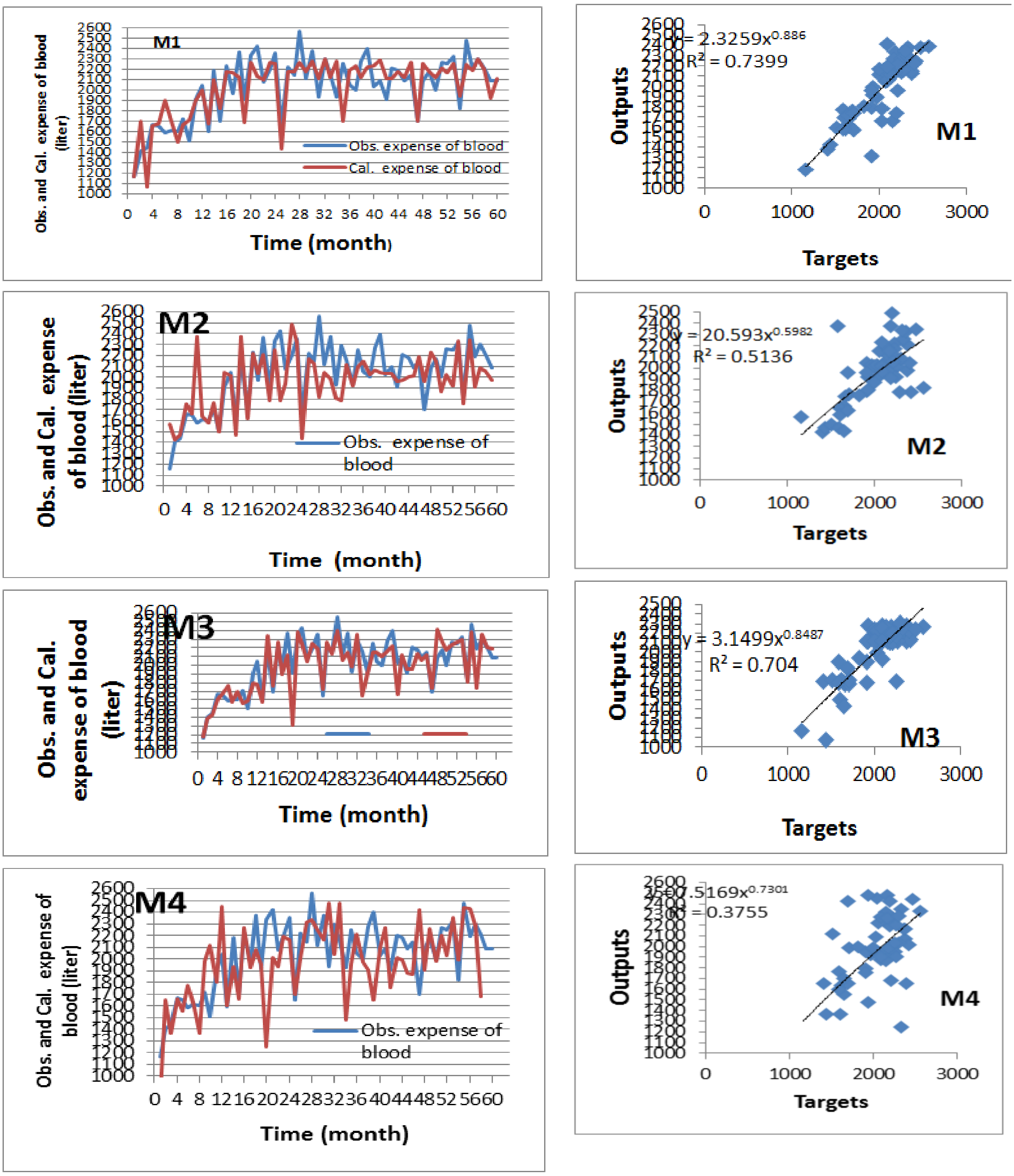

As evident from figure 2, the minimum error in expense of blood values has been obtained with 14 neurons in the hidden layer for all models. Table 2, shows the comparison of models computed over the test dataset, with marked values corresponding with best performance according to the criteria in each column. Model number 1 is the best in performance in testing stages compared to the other models, which has the lowest RMSE and MAE values with the highest R value.

**Table 2:** Performance parameters of the artificial neural network models in testing period with 14 nodes.

| Model Number | R | RMSE | MAE |
|---|---|---|---|
| 1 | 0.860174 | 108.8187 | 81.35443 |
| 2 | 0.716659 | 175.3731 | 148.8976 |
| 3 | 0839047 | 157.4312 | 102.0028 |
| 4 | 0.612781 | 262.519 | 195.8108 |

Application of artificial neural network models for prediction of expense of blood is demonstrated in figure 2, in the form of hydrograph and scatterplot. The figure also shows an analysis between the network outputs and the corresponding targets for the test dataset. Also, the figure shows the ability of artificial neural networks as a powerful tool to predict. The performance of the neural networks could be improved by using additional information related to the variable under consideration such as blood type.

**Figure 2**. Comparison of artificial neural network models for blood bank.




## Conclusions

In this study, an application of artificial neural network model was developed to predict the monthly expense of blood in a blood bank of Nasiriyah city. Data included monthly values of the expense of blood and the donors of blood for five years. The data was spilt into three data sets for training, validation and test in the ratio 70:15:15, respectively. Three layers feed the forward network with sigmoid hidden neurons and linear output neurons. The back-propagation algorithm gives a prescription for changing the weights in any feed forward network to learn a training vector of input-output pairs. The output explained that the artificial neural network with back-propagation algorithm is a powerful technique for predicting the expenses of blood and the best numbers of neuron in the hidden layer is equal to 14. Also, the results showed the efficiency of ANNs decreases when increasing the length of the forecasting period. The study illustrates practical application of ANNs approaches, adequately combined with other frequently used tools in the context of blood inventory systems planning and management. Hence, it can be concluded that the model is viable enough to be considered in future applications, competing with other classical techniques. The study recommends using hybrid systems developed from various artificial intelligence methods in order to get more accurate predictions.

**REFERENCES**

Ali, H., Ahmed, A. and Husham, T. (2017). Comparison of data-driven modelling techniques for predicting river flow in an arid region. International Journal of Applied Engineering Research, Vol. 12, No. 11, pp. 2647-2655.

Haddad, O. B., Sharife, F. and Alimohammadi, S. (2005). ANNs in river flow forecasting. Proceedings of the 6th WSEA Int. Conf. on Evolutionary Computing, Lisbon, Portugal, pp.316-324.

Haykin, S. (1999). Neural networks, a comprehensive foundation. Second edition. Prentice-Hall, New Jersey.

Heizer, J., Render, B. & Munson, C. (2017). Operations management: Sustainability and supply chain management.12th Ed, *New York, Pearson Education*.

Khairi, A. O. (2017). Prediction of dissolved oxygen in tigris river by water temperature and biological oxygen demand using artificial neural networks (ANNs). Journal of University of Duhok, Vol. 20, No. 1, pp. 691-700.

Krajewski, L. J., Ritzman, L. P. & Malhotra, M. K. (2013). Operations management: Processes and supply chains.10th Ed, New Jersey, Pearson Education.

Lippmann, R.P. (1987). An introduction to computing with neural nets.IEEE ASSP.

Masrur Ahmed, A. A. (2017). Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs). Journal of King Saud University-Engineering Sciences Vol. 29, No. 2, pp.151-158.

Nourani, V. and Babakhani, A. (2013). Integration of artificial neural networks with radial basis function interpolation in earth fill dam seepage modeling. Journal Comput Civil Engineering, Vol. 27, No. 2, pp.183-195.

Rankovic, V., Radulovi, J., Radojevic, I., Ostojic, A. and Comic, L. (2010). Neural network modeling of dissolved oxygen in Gruza reservoir, Serbia. Ecol,Model, Vol. 221, No. 8, pp.1239-1244.

Russell, R.S. & Taylor, B. W. (2011). Operations management creating value along the supply chain.7th Ed, John Wiley and Sons.

Schroeder, R. G. & Goldstein, S. M. (2018). Operations management in the supply chain: Decision and cases. 7th Ed, New York, McGraw-Hill Education.

Slack, N., Chambers, S. & Johnston, R. (2010). *Operations management*. Pearson education.6th Ed, London, Pearson Education Limited.

Schroeder, R. G. & Goldstein, S. M. (2018). Operations management in the supply chain: Decision and cases. 7th Ed, New York, McGraw-Hill Education.

Slack, N., Chambers, S. & Johnston, R. (2010). *Operations management*. Pearson education.6th Ed, London, Pearson Education Limited.

Schroeder, R. G. & Goldstein, S. M. (2018). Operations management in the supply chain: Decision and cases. 7th Ed, New York, McGraw-Hill Education.

Slack, N., Chambers, S. & Johnston, R. (2010). *Operations management*. Pearson education.6th Ed, London, Pearson Education Limited.

Stevenson, W. J. (2018). *Operations management* .13th Ed, *New York,* McGraw-Hill *Education*.

Suresh, N, & Kumar, S. A. (2008). Production and operations management (with skill development, caselets and cases).2th Ed, New Delhi, New Age International (P) Ltd.