# Utilizing Artificial Neural Networks for Predictive KPI Analysis in Bridge Projects

**Faiq M.S. Al-Zwainy[1*], Maryam G.S. Al-khazrajy[2], Nidal M. Hussein[3], Sherif Mohamed[4], Mazin M. Sarhan[5], Tariq J. Al-Musawi[6], Gasim Hayder[7]**

[1]Higher Institute of Forensic Sciences, , Al-Nahrain University,Baghdad, Iraq
[1]Forensic DNA Center for Research and Training, Al-Nahrain University, Baghdad, Iraq,
Email: faiqalzwainy@gmail.com
[2]Civil Engineering Department, Al-Nahrain University, Baghdad-Iraq,
Email: maryamdarwash97@gmail.com
[3]Department of Civil Engineering, Faculty of Engineering, University of Petra, Amman, Jordan,
Email: nidal.hussein@uop.edu.jo
[4]School of Surveying and Built Environment, University of Southern, Queensland, QLD, Australia,
Email: s.mohamed@griffith.edu.au
[5]Civil Engineering Department, College of Engineering, Mustansiriyah University, Baghdad, Iraq,
Email: mazin.m@uomustansiriyah.edu.iq
[6]Building and Construction Techniques Engineering Department, Al-Mustaqbal University College, 51001 Hillah, Babylon, Iraq, Email: tariq.jwad@uomus.edu.iq
70 Institute of Energy Infrastructure, Universiti Tenaga Nasional (UNITEN), Kajang 43000, Selangor Darul Ehsan, Malaysia
[7]Department of Civil Engineering, College of Engineering, Universiti Tenaga Nasional (UNITEN), Kajang 43000, Selangor Darul Ehsan, Malaysia, Email: gasim@uniten.edu.my
*Corresponding Author

Received: 10.07.2024 Revised: 16.08.2024 Accepted: 24.09.2024

**ABSTRACT**
KPI Analysis is a very important element in the process of construction project management especially with regard to the predictive of duration of bridge projects, main objective of this study is to develop intelligent prediction models, using Artificial Neural Network (ANN), to forecast earned value indicators in the early stages of the lifecycle of bridge projects in Republic of Iraq. Data used to develop neural network model for estimation of earned value indicators were past bridge contract data from Iraq, where, Iraqi Ministry of Construction, Housing, Municipalities, and Public Works launched bridges projects in 2021 and is expected to continue until 2025. variables affecting on earned value indicators of bridge projects were divided into two main categories; First category called Independent Variables, which include three variables, as the following; Schedule Performance Index (SPI), Cost Performance Index (CPI), and Cost-Schedule Index (CSI). While, second category called Dependent Variables, which include four variables as the following; Planned Cost (PC), Planned Duration (PD), Total Area (TA), and, Total Bridge Length (TBL). Three model was built for the prediction of earned value indectors of bridge projects. It was found that ANN models have the ability to predict SPI, CPI and CSI with excellent degree of accuracy of the coefficient of correlation (R) 0.8853%, 0.8037%, 0.81094%, and average accuracy percentage of 89.91%, 69.829%, 83.794% respectivily. This indicates that the relationship between the independent and independent variables of the developed models is good and the predicted values from a forecast model fit with the real-life data.

**Keywords**: Bridge Projects, Artuficial Neural Networks, KPI, SPI, CPI, CSI.

## 1. INTRODUCTION

Construction sector is one of the world's most complex, largest, and most challenging industries. Since the beginning of the 21st century, the Republic of Iraq has witnessed significant activity in the field of construction. Despite encompassing various variables and contradictions, the construction sector plays an active and influential role in implementing economic and social development plans globally. This sector is characterized by the diversity of its activities and events, represented by the different construction projects, each with unique circumstances and requirements. This imposes significant

burdens on the executing entities. The construction sector grows and expands over time with increased national income**(Fei, et al., 2021).**
Hence, the important role of project management emerges in completing construction projects within the specified time, required cost, and necessary quality**(Kerzner, 2017).**
One of the responsibilities facing this vital sector is overcoming failures in time and cost during the execution phases. Time and cost are critical factors for stakeholders, and all construction projects encounter numerous problems, dilemmas, complexities, and risks during the execution phase. These issues lead to the failure to complete construction projects within the time, cost, and quality specified during the planning phase. Since time and cost failures are among the reasons for project delays and loss of control over time and cost in construction projects, project management must implement effective measures to identify the causes of time and cost failures, adopt specific standards, and develop solutions and remedies to avoid and overcome them**(Mubarak, 2015)(Al-Zwainy et al, 2024).**
Accurate estimation in the early stages of the construction project lifecycle is a critical factor for its success. However, making a quick and precise estimate during the planning stage is challenging, especially if the contractor's documents, such as plans and quantities, still need to be completed. Therefore, various artificial intelligence techniques have been applied to achieve accurate estimation in the early stage when limited information about construction projects is available**(Zou, et al., 2017) (Al-Somaydaii et al, 2024).**
The main objective of this study is to develop intelligent prediction models, using Artificial Neural Network (ANN), to forecast earned value indicators in the early stages of the lifecycle of bridge projects in Republic of Iraq. To achieve this objective, it is necessary to identify the factors that affect the performance of bridge projects, which may be available in the early stages of the project lifecycle. Therefore, it attempts to develop and evaluate earned value indicator models through the following steps:

1) Collecting data and identifying variables for intelligent prediction models that affect earned value indicators for Iraqi bridge projects.
2) Developing mathematical models to forecast earned value indicators for Iraqi bridge projects.
3) Finding accurate mathematical equations to calculate earned value indicators for bridge projects.
4) Validating and verifying the accuracy of the artificial neural network models.

## 2. Data Collection and Identification of Variables

This study analyses historical data from bridge projects in the Republic of Iraq to collect and identify variables that influence the development of intelligent prediction models.
Historical data helps establish a relationship between the key factors affecting the earned value indicators of bridge projects, enabling accurate estimations for new projects**(Al-Marsomi and Al-Zwainy, 2023a)**.
Identifying and evaluating the factors that affect the earned value indicators of bridge projects is a critical issue faced by stakeholders in construction projects to control the performance of these projects. Understanding the key factors that influence earned value indicators, positively or negatively, can contribute to developing a strategy to mitigate deficiencies and enhance the effectiveness of bridge projects**(Khan et al, 2023)**. Therefore, there is an urgent need to identify and understand the various factors affecting earned value indicators to focus on the necessary steps to reduce deviations in time, cost, or delays in completing bridge projects, thereby increasing productivity and overall project performance**(Hussein and Al-Zwainy, 2024).**
This study describes the development of neural network models to forecast earned value indicators of bridges projects based on recent historical projects data. The initial impetus was paucity of data available that can provide reliable information about earned value indicators. Data used to develop neural network model for estimation of earned value indicators were past bridge contract data from Iraq, where, Iraqi Ministry of Construction, Housing, Municipalities, and Public Works launched bridges projects in 2021 and is expected to continue until 2025. These projects aims to improve the country's transportation infrastructure and facilitate traffic movement between cities and regions.
First package of bridge projects in Iraq includes thirty projects with a total cost of 1.634 billion Iraqi dinars and is scheduled to be completed within three years. These projects involve the construction of new bridges and the improvement of existing ones. The goals of these projects are to enhance traffic flow and alleviate traffic congestion in Iraq, create new job opportunities for Iraqis during the construction and operation phases, improve the standard of living by facilitating access to essential services and economic opportunities, and stimulate the economy by attracting investments and boosting trade**(Al-Zwainy and Al-Marsomi, 2023b)**.
Data collection method used in this study is the direct data gathering from bridge Construction Companies and the direct interview with the concerned managers and engineers. This method faces a great difficulty nowadays because of the shortage in documentation. In spite of these obstacles, the

researcher succeeded in gathering well trusted data for more than thirty projects through the companies' visits and reading the concerned sheets, documents and reports for bridge projects**(Risan et al, 2024).**
In this study, factors or variables affecting on earned value indicators of bridge projects were divided into two main categories; First category called Independent Variables, which include three variables, as the following; Schedule Performance Index (SPI), Cost Performance Index (CPI), and Cost-Schedule Index (CSI). While, second category called Dependent Variables, which include four variables as the following; Planned Cost (PC), Planned Duration (PD), Total Area (TA), and, Total BridgeLength (TBL).
Each one of the variables has been analysed in order to determine the best way of representation in the modelling process. The way in which these variables are represented are real numbers. Table (1) illustrates the summary of statistical values for the independent and dependent variables for thirty bridge projects in the Republic of Iraq, influencing the dependent variables of the earned value indicators used in artificial neural network models.

**Table 1:** Statistical Values ofIndependent and Dependent Variables.

| **Statistical Values** | **Inputs (Independent Variables)** | | | |
|---|---|---|---|---|
| | Planned Cost (PC) (ID) | Planned Duration(PD) (Day) | Total Area (TA) (M2) | Total Bridge Length (M.L) |
| **Max.** | 220000000000 | 580 | 70000 | 2280 |
| **Min.** | 80325000000 | 240 | 24100 | 1000 |
| **Range** | 139675000000 | 340 | 45900 | 1280 |
| **Average** | 150162500000 | 410 | 47050 | 1640 |
| **Standard Deviation** | 69837500000 | 170 | 22950 | 640 |
| | **Outputs (Dependent Variables)** | | | |
| | Scheduling Performance Index (SPI) | CostPerformance Index (CPI) | Cost-Schedule Index (CSI) | |
| **Max.** | 1.25 | 0.50 | 0.39 | |
| **Min.** | 0.78 | 0.17 | 0.19 | |
| **Range** | 0.47 | 0.33 | 0.20 | |
| **Average** | 1.015 | 0.335 | 0.29 | |
| **Standard Deviation** | 0.235 | 0.165 | 0.1 | |

### 3. Developing KPI$_S$-ANN Models

Methodology used in developing mathematical models to predict earned value indicators for Iraqi bridge projects in this current study involves three main phases:
**1) Fist Stage**: At the conceptual analysis stage, a Neural Network paradigm has to be selected as a suitable environment for developing the application. It can be done based on a comparison of the application requirements against neural network paradigm capabilities. Based on the literature review of Neural Network, the neural network type deemed suitable for cost estimation has been identified as feed-forward pattern- recognition type (Back propagation) to suit the desired interpolative and predictive performance of the model**(Hammoody et al, 2022)**.
**2) Second Stage**: Problem analysis is the identification of the independent factor(s) that fully describes the problem and that are expected to be easily obtainable for training the neural network**(Nabeel and Faiq, 2017)**. For this present study, four major factors describing a bridges projects and affecting its earned value indicators have been identified. These factors include descriptors of Planned Cost (PC), Planned Duration (PD), Total Area (TA), and, Total Bridge Length (TBL).so as to have a generic estimating model accounting for time, capacity and other uncertainty-related factors.
**3) Third Stage**: Problem structuring, entails the arrangement and representation of the descriptive factors and their associated results in the form of inputs and outputs, as required by neural network modelling**(Raheem and Al-Zwainy, 2020)**. Four inputs readily were identified, therefore, outputs describing earned value indicators of a bridges projects can be modelled in different ways and thus one neural networks model were constructed with four inputs and three outputs are Schedule Performance Index (SPI), Cost Performance Index (CPI), and Cost-Schedule Index (CSI).
**4) Fourth Stage**: Implementation of Neural Network model: In this study, can be used a computer program called Neuframe to build and develop the neural network model. Neuframe software: This

commercial program, written in C++ and classified under artificial intelligence technologies, is easy to use for building various artificial neural networks**(Al-Zwainy et al, 2024)**. It was used to illustrate the process of constructing sub-models, starting from the input model and ending with the output model, as shown in Figure (1). To complete the ANN systems, the individual Neuframe components must be explained as shown in Table (2).

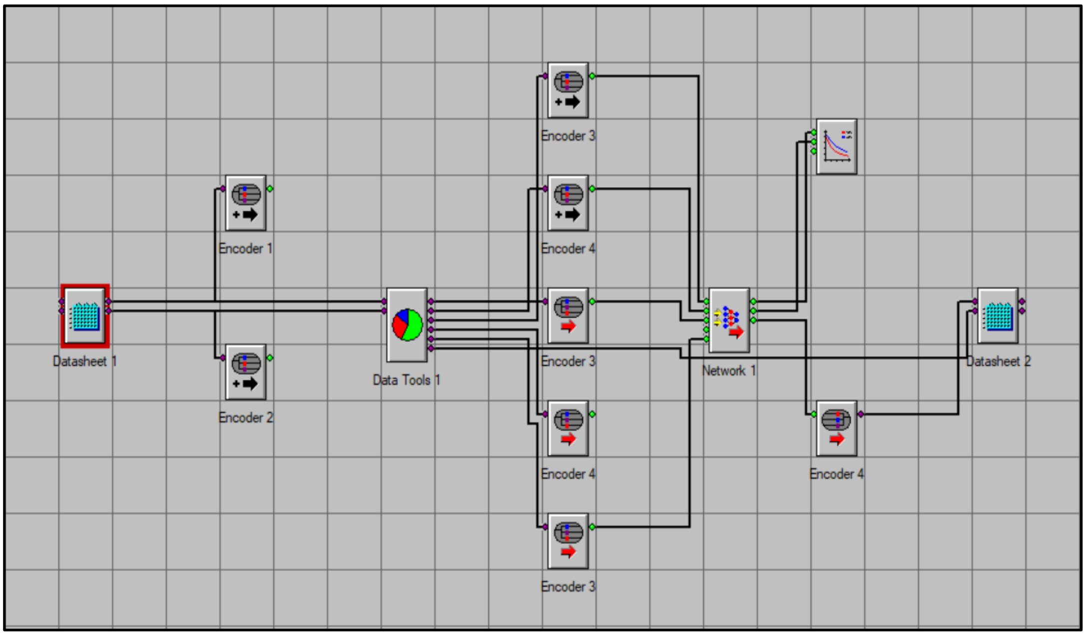
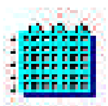
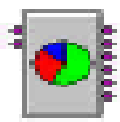
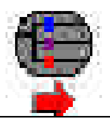
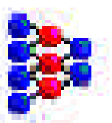



**Figure 1:**Architecture of Neuframe software

**Table 2:** Components of Neuframe software.

| | |
|---|---|
| **Data Sheet**  | The datasheet provides the storage of data within the workspace. Data may be imported from a range of popular formats into the workspace main table. From the main table, different subsets of the data called Data Views can be created, and then used by other objects.A complementary set of export formats allow for the easy transfer of data into other programs.The datasheet supports the familiar user functions found in other Windows spreadsheets such as right mouse button menus, cut, copy and paste, user definable column headings etc. |
| **Data Tools**  | This valuable addition to Issue 4.0 of Neuframe takes of lot of the tedious data selection out of the set-up process. A data tool automatically takes a complete data set and splits it into Training Data, Test Data and Query Data. This timesaving tool also handles all of the preparation necessary to set up a Time Series data table, with a high degree of flexibility in selecting the time series parameters. |
| Encoding  | This provides a more comprehensive set of pre and post processing functions for both textual and numerical information. The Encoder prepares data for presentation to an object such as a neural net and also takes the output from the net and decodes the results back to meaningful information. |
| **Standard Back Propagation (SBP) Network**  | This is a fast feed forward net using back propagation for training. One of the most versatile and consistent of the variety of neural nets available, it is used to model patterns within data and therefore has wide and extensive application across many areas, including decision support, process modeling and diagnosis. A neural network is automatically created based on the data to be used. For the expert user, extensive set up options allow a fine control over the design of the net, from the number of layers, to the transfer function used in each layer. |

| | |
|---|---|
| **Error Graph Tool**  | This object may be connected to nets to show the progress of training and validation of the SBP algorithm. The tool provides a valuable visual feedback of progress. Root Mean Squared (RMS) error of the training data and test data are displayed against the training epoch (one training pass). The graph is dynamic and plots training progress as it occurs. |

**5) Five Stage**: methodology of KPIs-ANN Models: The methodology used in developing the artificial neural network model in this study involves developing a set of sub-models, such as the Input Model, Output Model, Data Division Model, Network Architecture Model, Weight Model, Learning Rate Model, Momentum Term Model, Transfer Function Model, Mathematical Equations Model, and Validation Model, as following;

**a) Inputs and Outputs of KPI$_S$-ANN Model**

Identifying and selecting variables in input and output models is critical for enhancing the artificial neural network's performance. Increasing the number of input and output variables significantly impacts the expansion of the artificial neural network's size, leading to improved efficiency. Current study employed the prior knowledge approach to determine the number of variables in the input and output models, this method is widely used in project management and validated in many research studies and scientific literature. Therefore, the input model included the independent variables, which are; Planned Cost (PC), Planned Duration (PD), Total Area (TA), and, Total Bridge Length (TBL). While, output model included the dependent variables, which are; Schedule Performance Index (SPI), Cost Performance Index (CPI), andCost-Schedule Index (CSI).

**b) Data Division of KPI$_S$-ANN Model**

Input or output data in the artificial neural network can be divided into three main groups; Training Set (TrS), Testing Set (TeS), and, Validation Set (VaS). TrS optimizes the weights associated with the artificial neural network, while VaS evaluates the network's performance at various training stages. Finally, TeS assesses the model's performance, therefore, dividing the data into the three categories mentioned above is a critical and fundamental step in developing artificial neural networks.

In current study, data was divided into training, testing, and validation sets using a difficulty and error method, the researcher adjusted the proportions of data in these sets to improve the performance of the artificial neural network. The goal was to achieve the highest correlation coefficient value, reflecting the strength and quality of the relationship between the predicted and actual earned value indicators, with minimize testing error rate. This study selected the best data division based on these two criteria, resulting in a split of 60% for Training Set (TrS), 20% for Testing Set (TeS), and 20% for Validation Set (VaS), as shown in Figure (2).

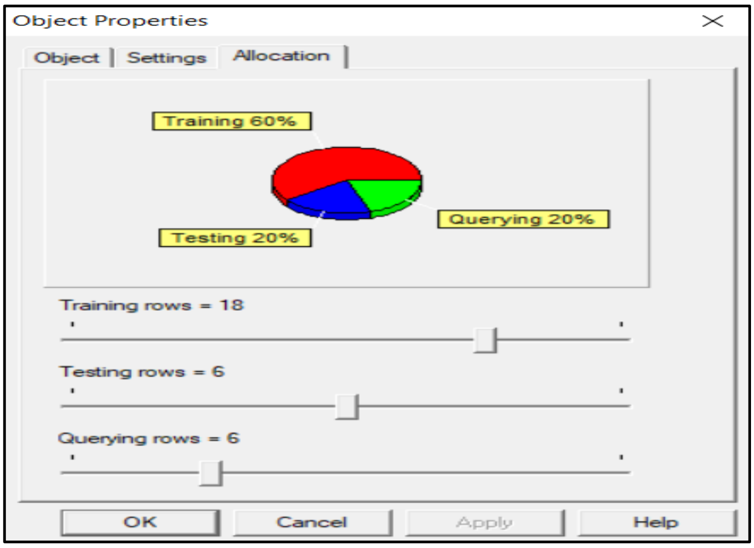



**Figure 2:** Data Distribution for Training, Testing, and Validation Sets.

Table (3) shows the data distribution percentages for training, testing, and validation sets. training set contains 18 samples at 60.0%, testing set contains 6 samples at 20.0%, and validation (Querying rows) set includes 6 samples at 20.0%.Total number of samples is 30, representing 100%. These samples were divided to improve the model's performance and ensure the accuracy of its predictions.

**Table 3:** Data Ratio for Training, Testing, and Investigation Set**.**

| Groups | | No. | Percent |
|---|---|---|---|
| **Sample** | Training Group | 18 | 60.0% |
| | Testing Group | 6 | 20.0% |
| | Investigation group | 6 | 20.0% |
| Valid | | 100 | 100.0% |
| Total | | 100 | 100.0% |

To distribute the total data of the independent and dependent variables, consisting of 30 projects, into three groupstraining, testing, and validationthree methods were used; Random Method: Neuframe software distributes variable data among three groups using a random distribution.Strip Method: This method divides the entire dataset into random subsets, each containing data for the training, testing, and validation groups.Blocked Method: This method treats the entire dataset as a single batch and sequentially divides it into three groups.
Random division method was used, as shown in Figure (3), where data was distributed with 60% for the training set, 20% for the testing set, and 20% for the validation set. This method achieved the lowest testing error rate of (0.212804) and a training error rate of (0.245944) as shown in Figure (4).

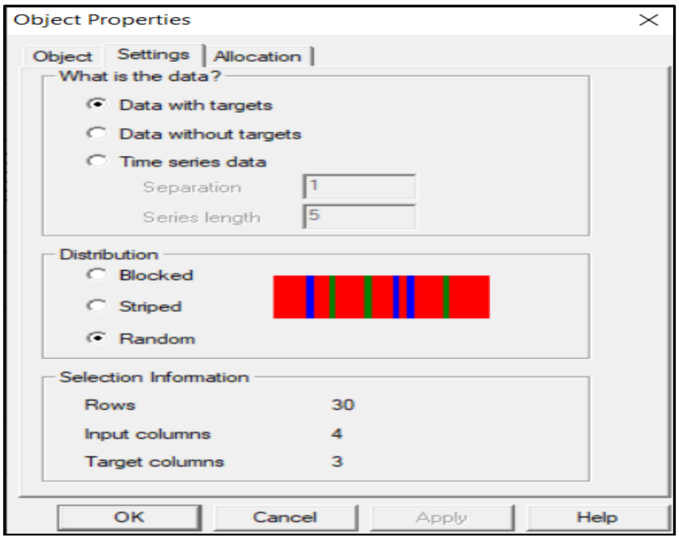



**Figure 3:** Data Division Using the Random Distribution Method

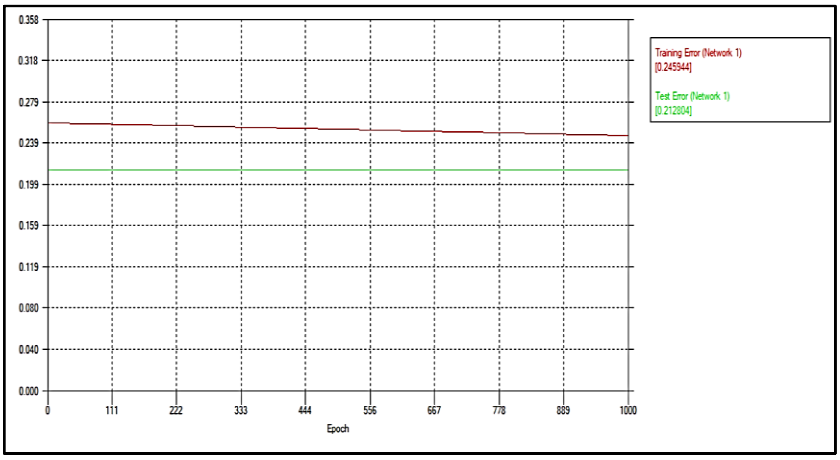



**Figure 4:** Training and Testing Error of $KPI_S$-ANN Model

Once the available data have been divided into their subsets, input and output variables are pre-processed by scaling them to eliminate their dimension and to ensure that all variables receive equal attention during training. Scaling has to be commensurate with the limits of the transfer functions used in the hidden and output layers (i.e. –1.0 to 1.0 for tanh transfer function and 0.0 to 1.0 for sigmoid transfer function). The simple linear mapping of the variables, extremes to the neural network's practical extremes is adopted for scaling, as it is the most commonly used method, **(Jaber et al, 2019)(Zamim et al, 2019).** As part of this method, for each variable x with minimum and maximum values of X min and X max, respectively, the scaled value xn is calculated as follows:

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad \text{.......................................................} (1)$$

**c)Hidden Layer of KPI$_S$-ANN Model**

Architecture of the artificial neural network relates to how the neurons are connected to form the network. Determining the optimal number of neurons in the hidden layer of the artificial neural network is a critical factor in its effectiveness. The number of neurons in the input layer is proportional to the number of elements that affect the calculation of earned value indicators. These elements include the planned value, earned value, and actual cost, which means they encompass three factors. output layer contains three neurons, which correspond to Schedule Performance Index (SPI), Cost Performance Index (CPI), and Cost-Schedule Index (CSI).

Multiple techniques exist to determine the optimal number of neurons in neural networks, and the best option is to use Equation (2). This involves starting with one neuron in the hidden layer and gradually increasing the number of neurons until optimal neural network performance is achieved **(Aidan et al, 2020).** In this study, adopted this approach in the study.

Number of nodes = (number of inputs * 2) + 1 .........................(2)

Table (4) clearly shows a slight difference in the error rate among the testing set. network performs best when the number of neurons is set to one, as evidenced by the highest correlation coefficient (99.96%) and the lowest testing error rate (0.212804%). Therefore, most efficient structure for this network, as determined in this study, consists of three layers: an input layer, a hidden layer, and an output layer. input layer of the neural network does not participate in any calculations. Its primary function is to receive the database and send it to the network. input layer passes the information to the hidden layer, which transmits the information to the output layer. Data is processed only within the hidden layer and the output layer.

**Table 4:**Effect of Number of Hidden Nodes on KPIS-ANN Model

| **No. of Nodes in hidden layer** | **Training Error %** | **Testing Error %** | **Coefficient Correlation(r)%** |
|---|---|---|---|
| 1 | 0.198725 | 0.212804 | 99.96 |
| 2 | 0.199925 | 0.254309 | 96.91 |
| 3 | 0.209988 | 0.266301 | 95.88 |
| 4 | 0.209988 | 0.266301 | 95.88 |
| 5 | 0.209988 | 0.266301 | 95.88 |
| 6 | 0.209988 | 0.266301 | 95.88 |
| 7 | 0.209988 | 0.266301 | 95.88 |
| 8 | 0.209988 | 0.266301 | 95.88 |
| 9 | 0.209988 | 0.266301 | 95.88 |

**d) Momentum Term and Learning Rate of KPI$_S$-ANN Model**

These two models play a vital role in enhancing the efficiency of the artificial neural network, working together within the network design, with each influencing the other. To evaluate the impact of the momentum term on the efficiency of the artificial neural network, the researcher experimented on the neural network, setting the momentum value at 0.8 and the learning rate at 0.2, as shown in Figure (5). Results showed that these values achieved optimal network performance, with lowest testing error rate (0.212804%) and highest correlation coefficient (99.96%). Neural network model performance remains stable and effective at these specified values, reflecting a good balance between learning speed and model stability.

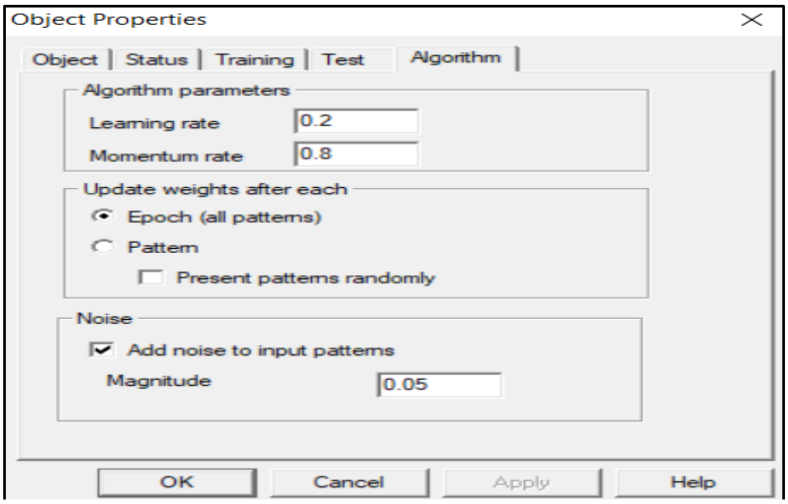


**Figure 5:**KPIS-ANN Model Parameters: Learning Rate and Momentum Term

**e) Transfer Function of KPI$_S$-ANN Model**

A primary experiment was conducted to study the impact of the activation function, using the sigmoid activation function for both the hidden and output layers. The researcher found that the sigmoid activation function is the primary activation function for neurons due to its ability to introduce non-linear manners into the computations of artificial neural networks by transforming the activation value to a range between 0 and 1. Moreover, it provides the additional advantage of having a simple derivative, which is essential in the backpropagation algorithm, a supervised learning technique used in feedforward networks in this study.

**f)Architecture of KPI$_S$-ANN Model**

Final design of the artificial neural network consists of three layers (input layer, output layer, and hidden layer) connected in a single loop, as shown in Figure (6).

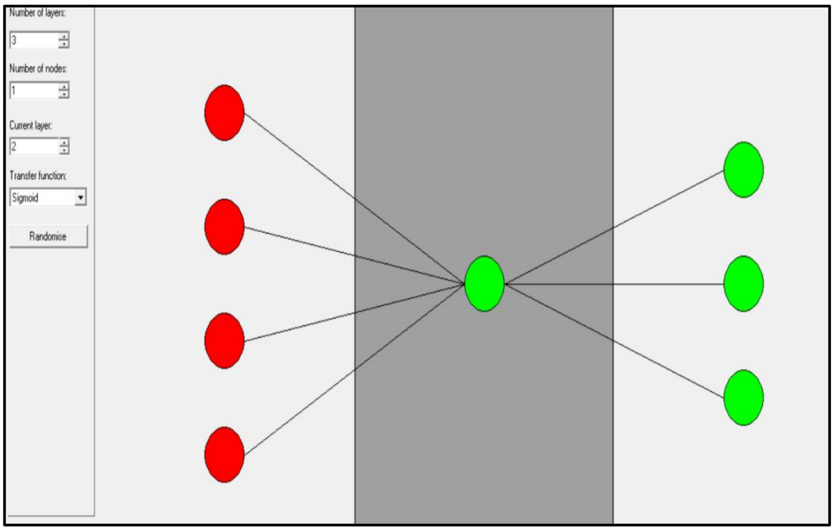


**Figure 6:** Architecture of KPI$_S$-ANN

**g) Weight of KPI$_S$-ANN Model**

Relationship between neurons in an artificial neural network is determined by the weight assigned to each connection. This value measures the significance of the connection between the two neurons. Each neuron multiplies every input value it receives from the neurons in previous layer by corresponding weights of the connections. The neuron then sums the products and adds a threshold value. After that, result is subjected to a transfer function that varies based on the type of neuron. output of the transfer function is then passed as the neuron's output to the neurons in the next layer**(Al-Zwainy et al, 2018)**. After completing training of artificial neural network, weight values for the connections between the input layer and the hidden layer, as well as weights between hidden layer and output layer, were obtained, as shown in Table (5).

**Table 5:** Weights and Threshold Levels for KPIS-ANN Optimal Model

| **Hidden layer nodes** | **$w_{ji}$ (weight from node i in the input layer to node j in the hidden layer)** | | | | **Hidden layer threshold $\theta_j$** |
|---|---|---|---|---|---|
| | **$W_{5-1}$** | **$W_{5-2}$** | **$W_{5-3}$** | **$W_{5-4}$** | |
| **j=5** | **1.12** | **-0.70** | **-4.39** | **0.91** | **3.69** |
| **Output layer nodes** | **$w_{ji}$ (weight from node i in the hidden layer to node j in the output layer)** | | | | **Output layer threshold $\theta_j$** |
| | **$W_{6-5}$** | **$W_{7-5}$** | **$W_{8-5}$** | | |
| **j=6** | **2.17** | **-** | **-** | | **-2.15** |
| **j=7** | **-** | **-3.17** | **-** | | **2.69** |
| **j=8** | **-** | **-** | **-4.93** | | **3.94** |
| **Parameters Effect** | • Model No. ($KPI_S$-ANN-1),<br>• Choices of division data (Striped)<br>• No. of Hidden Nodes (1)<br>• Momentum Term (0.8)<br>• Learning Rate (0.2)<br>• Activation Function (Sigmoid) | | | | |

**h) Equation of $KPI_S$-ANN Model**

Using the connection weights and the threshold levels shown in Table (5), prediction equations of $KPI_S$-ANN Model can be expressed as follows:

1) Equation of ANN-SPI Model:

$$SPI=\frac{\text{Range (SPI)}}{1+e^{-x_2}}+\text{Min. (SPI)} \quad \text{.......(3)}$$

$$SPI=\frac{0.47}{1+e^{-x_2}}+0.78 \quad \text{.......(4)}$$

$$x_1=[(PD*wight)+(PC*weight)+(TBL*weight)+(TA*weight)]+\theta \quad \text{.......(5)}$$

$$x_1=[(PD*1.12)+(PC*-0.7)+(TBL*-4.39)+(TA*0.91)]+3.69 \quad \text{.......(6)}$$

$$x_2=(X_1*Weight)+\theta \quad \text{.......(7)}$$

$$x_2=(X_1*2.17)+(-2.15) \quad \text{.......(8)}$$

$$SPI=\frac{0.47}{1+e^{-(-2.15+2.17X_1)}}+0.78 \quad \text{.......(9)}$$

2) Equation of ANN-CPI Model

$$CPI=\frac{\text{Range (CPI)}}{1+e^{-x_2}}+\text{Min(CPI)} \quad \text{.......(10)}$$

$$CPI=\frac{0.33}{1+e^{-x_2}}+0.17 \quad \text{.......(11)}$$

$$x_1=[(PD*1.12)+(PC*-0.7)+(TBL*-4.39)+(TA*0.91)]+3.69 \quad \text{.......(12)} \quad x_2=(X_1*Weight)+\theta \quad \text{.......(13)}$$

$$CPI=\frac{0.33}{1+e^{-(-2.15+2.17X_1)}}+0.17 \quad \text{.......(15)}$$

3) Equation of ANN-CSI Model

$$CSI=\frac{\text{Range (CSI)}}{1+e^{-x_2}}+\text{Min(CSI)} \quad \text{.......(16)}$$

$$CSI=\frac{0.20}{1+e^{-x_2}}+0.19 \quad \text{.......(7)}$$

$$x_1=[(PD*1.12)+(PC*-0.7)+(TBL*-4.39)+(TA*0.91)]+3.69 \quad \text{.......(18)}$$

$$x_2=(X_1*Weight)+\theta \quad \text{....... (19)}$$

$$x_2=(X_1*2.17)+(-2.15) \quad \text{....... (20)}$$

$$CSI=\frac{0.20}{1+e^{-(-2.15+2.17X_1)}}+0.19 \quad \text{....... (21)}$$

**i) Verification and Validationof $KPI_S$-ANN Model.**

Verification and validation of an Artificial Neural Network (ANN) model are crucial steps in ensuring its accuracy and reliability. Verification involves checking if the ANN model is implemented correctly and if it

produces the expected results. Validation involves checking if the ANN model is accurate and reliable on unseen data (Jasimet al ,2020)(Jaber et al, 2019).
Four statistical criteria can be used to demonstrate the accuracy and effectiveness of the mathematical equations derived from the artificial neural network models for determining the Schedule Performance Index (SPI), Cost Performance Index (CPI), and Cost-Schedule Index (SCI). Statistical criteria used in development KPIS-ANN Models include Mean Absolute Percentage Error (MAPE), Average Accuracy Percentage (AA%), Coefficient of Determination ($R^2$), and Correlation Coefficient (R), are as follows:

1) Mean Absolute Percentage Error (MAPE)
   a. MAPE = (∑(│A-E│)/A * 100%)/n………………… (22)
2) Average Accuracy Percentage (AA%)
   a. AA% = 100% - MAPE ……………………..……… (23)
3) Coefficient of Determination ($R^2$).
4) Coefficient of Correlation (R).

Where:
A: Actual values
E: Estimtion values
n: Number of projects
Tables (6), (7), and (8) represent the results of the four statistical criteria mentioned above for many projects (six projects), which represent (20%) of the total data fed into the software (neuframe) used. After building the three KPIS-ANN models, these were classified into the validation group.

**First: ANN-SPI Model**
From Table (6) and Figure (7), it can be concluded that the Schedule Performance Index (SPI) model has outstanding performance, as it has a high correlation (R) of 88.53%) and a coefficient of determination ($R^2$) of (78.39%). The model's accuracy (AA%) is (89.91%). Therefore, it can be inferred that the SPI model shows excellent agreement with the actual values.

**Table 6:** Verification and Validation of ANN-SPI Model

| **Projects** | **Actual SPI** | **Estimate SPI** | **MAPE%** |
|---|---|---|---|
| 1 | 0.94 | 0.95 | 1.060 |
| 2 | 0.78 | 0.97 | 24.36 |
| 3 | 1.25 | 1.20 | 4.000 |
| 4 | 1.00 | 0.97 | 3.000 |
| 5 | 0.78 | 0.93 | 19.23 |
| 6 | 0.90 | 0.98 | 8.890 |
| **Correlation Coefficient (R)** | 88.53% | | 60.54/6= 10.090 |
| **Coefficient of Determination ($R^2$)** | 78.39% | | |
| **Average Accuracy (AA%)** | 100%-10.090=89.91% | | |

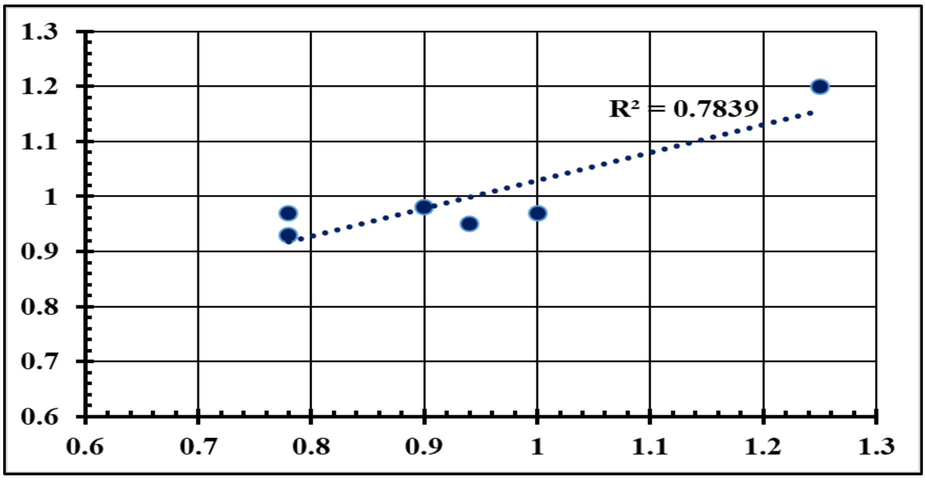


**Figure 7:**Validation of ANN-SPI Model

The following statistical tests were conducted on “R” (the coefficient of correlation) value for ANN-SPI Model, where $R^2$=0.7839, N=150:

1) Probable Error (P.E.) in “R” value

$$P.E. = 0.6745\left[\frac{\left(1-R^2\right)}{\sqrt{N}}\right] \quad \text{.............................................. (24)}$$

P.E. = 0.05950 therefore, R=0.8853 ± 0.05950

According to (Hussam et al, 2024); the probable error is regarded as a measure of significance of Karl Person’s coefficient of correlation (R), and if the probable error is small (compared with R), correlation directly exists where R (0.9448)> 0.5; Hence, the correlation of the studied cost equation is existing.

2) Standard Error (S.E.) in “R” value

$$S.E. = \left(\frac{1+R^2}{\sqrt{N}}\right) \quad \text{..................................................... (25)}$$

S.E. = 0.728271, The correlation is accepted for R=0.8853, and 6 observations.

3) Test of significance

(Hussam et al, 2024), indicates that the correlation may be accepted when R>0.22 (for 6 observations). Again, the correlation is accepted for R=0.8853, and 6 observations.

4) A simple method of testing whether “R” differs significantly from “zero”.

Taking a null hypothesis that there is no correlation between the two variables, provided “N” is large:

$$\frac{3}{\sqrt{N}} \quad \text{.............................................................. (26)}$$

IF the value arrived at by this test is greater than the observed or computed value of correlation coefficient (R≤ $\frac{3}{\sqrt{N}}$ ), the correlation is not significant. (Hussam et al, 2024);

$$\frac{3}{\sqrt{6}} = 1.2247 \geq 0.8853$$

Hence, coefficient of correlation can be taken as significant.

5) “t” test

There is another test of significance of coefficient of correlation, in which the value of “t” is computed by the following formula:

$$t = \frac{R*\sqrt{N-2}}{\sqrt{1-R^2}} \quad \text{...................................................... (27)}$$

IF the computed value of “t” is greater than the table value, the correlation is taken as significant.

$$t = \frac{R*\sqrt{N-2}}{\sqrt{1-R^2}} = t = \frac{0.8853*\sqrt{6-2}}{\sqrt{1-0.7839}}$$ =3.80884> tabulated “t”.

This means that correlation coefficient is highly significant.

Finally, it can be seen that the ANN-SPI Model was accurate in terms of MAPE, AA, R and $R^2$.

**Second: ANN-CPI Model**

Based on the data presented in Table (7), it can be concluded that the Cost Performance Index (CPI) model has outstanding performance. This is evident from its strong correlation (R) of (80.37%) and a coefficient of determination ($R^2$) of (64.59%). The model's accuracy is (69.829%). Therefore, it can be inferred that the CPI model shows exceptional concurrence with the actual data, as illustrated in Figure (8).

**Table 7:** Verification and Validation of ANN-CPI Model.

| **Projects** | **Actual CPI** | **Estimate CPI** | **MAPE%** |
|---|---|---|---|
| 1 | 0.22 | 0.35 | 59.09 |
| 2 | 0.29 | 0.33 | 13.79 |

| | | | |
|---|---|---|---|
| 3 | 0.17 | 0.31 | 82.35 |
| 4 | 0.38 | 0.33 | 13.15 |
| 5 | 0.50 | 0.55 | 10.00 |
| 6 | 0.38 | 0.39 | 2.630 |
| **Correlation Coefficient (R)** | 80. 37% | | 181.0264/6= 30.171 |
| **Coefficient of Determination ($R^2$)** | 64.59 % | | |
| **Average Accuracy (AA%)** | 100%-30.171=69.829% | | |

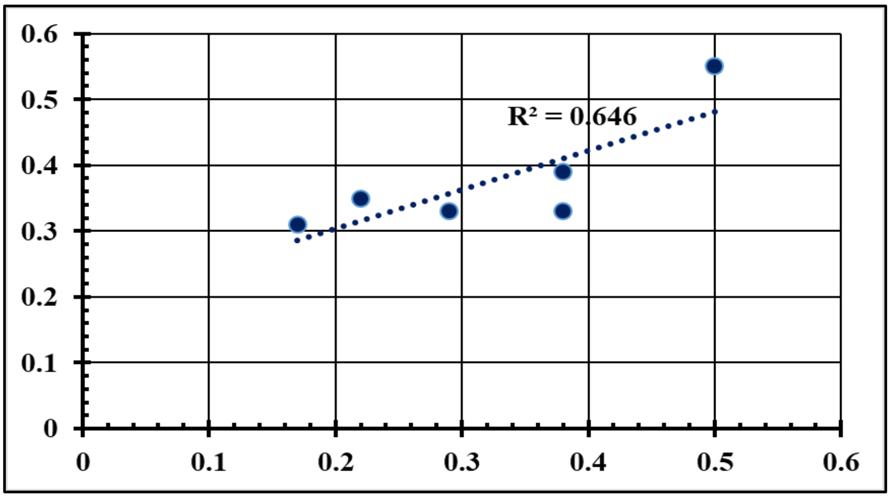



**Figure 8:** Validation of ANN-CPI Model

The procedure of the statistical test on “R” value for ANN-CPI Model is the same as that used in the development of ANN-SPI Model, as shown in Table (8), where R=0.8073, $R^2$=0.646

**Table 8:** Results of Statistical Test on "R" Value for ANN-CPI Model

| **Statistical Tests** | **Value** | **Result** |
|---|---|---|
| Probable Error P.E. | 0.097478 | Correlation is existing, since R=0.8037+0.097478>0.5 |
| Standard Error S.E. | 0.67197387 | correlation is accepted |
| Test of significance | R >0.22 | correlation is accepted |
| simple method of testing | R<1.2247 | correlation can be taken as significant |
| “t” test | 2.70161> t- tabulated | correlation is highly significant |

Finally, it can be seen that the ANN-CPI Model was accurate in terms of MAPE, AA, R and $R^2$.

**Third: ANN-CSI Model**

Based on the data presented in Table (9), it can be concluded that ANN-CSI model has exceptional performance. This is evident from its strong correlation (R) of 81.094% and a coefficient of determination ($R^2$) of (65.76%). model's accuracy, expressed as (AA%) is (83.794%). Therefore, it can be inferred that the ANN-CSI model shows high concurrence with the actual data, as illustrated in Figure (9).

**Table 9:** Verification and Validation of ANN-CSI Model.

| **Projects** | **Actual CSI** | **Estimate CSI** | **MAPE%** |
|---|---|---|---|
| 1 | 0.22 | 0.29 | 31.81 |
| 2 | 0.25 | 0.27 | 8.000 |
| 3 | 0.25 | 0.25 | 0.000 |
| 4 | 0.38 | 0.39 | 2.630 |
| 5 | 0.39 | 0.32 | 17.94 |

| | | | |
|---|---|---|---|
| 6 | 0.19 | 0.26 | 36.84 |
| **Correlation Coefficient (R)** | 81.094% | | 97.2405/6= 16.206 |
| **Coefficient of Determination ($R^2$)** | 65.76% | | |
| **Average Accuracy (AA%)** | 100%-16.206=83.794% | | |

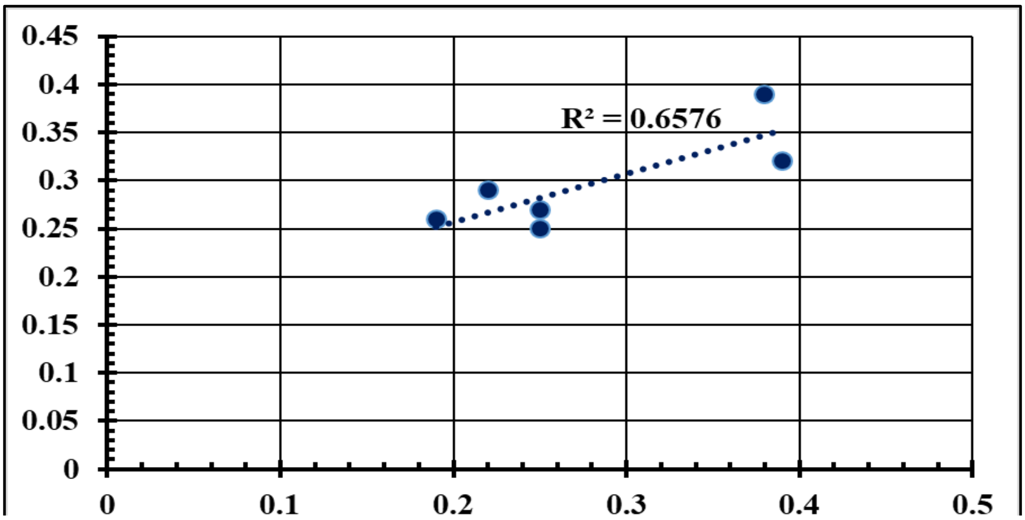



**Figure 9:** Validation of ANN-CSI Model.

The procedure of the statistical test on "R" value for ANN-CSI Model is the same as that used in the development of ANN-SPI Model and ANN-CPI Model, as shown in Table (10), where R=0.81094, $R^2$=0.6576

**Table 10:** Results of Statistical Test on "R" Value for ANN-CSI Model

| Statistical Tests | Value | Result |
|---|---|---|
| Probable Error P.E. | 0.094284 | Correlation is existing, since R=0.81094+0.094284>0.5 |
| Standard Error S.E. | 0.676709 | correlation is accepted |
| Test of significance | 0.22>R | correlation is accepted |
| simple method of testing | R<1.2247 | correlation can be taken as significant |
| "t" test | 2.771738> t- tabulated | correlation is highly significant |

Finally, it can be seen that the ANN-CSI Model was accurate in terms of MAPE, AA, R and $R^2$.

## 4. CONCLUSIONS

KPIS-ANN Model would make a useful benchmark against which neural network models could be measured in bridges projects, Therefore, a high predication accuracy requires much time to train network and search for sophistication model. Table (411) represents the results of the four statistical criteria mentioned earlier for six bridge projects. These projects represent the percentage of the total data entered into the program and were classified into the validation group after the completion of creating the artificial neural network model.

**Table 11:** Results of KPIS-ANN Model Investigation.

| $KPI_S$-ANN Model | (R) | ($R^2$) | MAPE% | (AA%) |
|---|---|---|---|---|
| **ANN-SPI Model** | 0.8853 | 0.7839 | 10.090 | 89.91 |
| **ANN-CPI Model** | 0.8037 | 0.6459 | 30.171 | 69.829 |
| **ANN-CSI Model** | 0.81094 | 0.6576 | 16.206 | 83.794 |

**REFERENCES**


[1] Aidan, I. A., Al-Jeznawi, D., & Al-Zwainy, F. M. (2020). Predicting earned value indexes in residential complexes' construction projects using artificial neural network model. International Journal of Intelligent Engineering and Systems, 13(4), 248-259. DOI: 10.22266/ijies2020.0831.22

[2] Al-Marsomi, M. S. K., & Al-Zwainy, F. M. S. (2023a). Assessing obstacles in construction-phases for regional development programs RDPs. Asian Journal of Civil Engineering, 24(8), 3425-3436.https://doi.org/10.1007/s42107-023-00723-0

[3] Al-Somaydaii, J. A., Albadri, A. T., & Al-Zwainy, F. M. (2024). Hybrid approach for cost estimation of sustainable building projects using artificial neural networks. Open Engineering, 14(1), 20220485.https://doi.org/10.1515/eng-2022-0485

[4] Al-Zwainy, F. M., Zaki, R. I., Al-saadi, A. M., & Ibraheem, H. F. (2018). Validity of artificial neural modeling to estimate time-dependent deflection of reinforced concrete beams. Cogent Engineering, 5(1), 1477485.https://doi.org/10.1080/23311916.2018.1477485

[5] Al-Zwainy, F., & Al-Marsomi, M. (2023b). Structural equation modeling of critical success factors in the programs of development regional. Journal of Project Management, 8(2), 119-132.DOI: 10.5267/j.jpm.2022.11.002

[6] Al-Zwainy, F., Abdalkarim, E. K., Majeed, W. K., Huseen, E. S., & Jari, H. S. (2024). Development Artificial Neural Network (ANN) computing model to analyses men's 100-meter sprint performance trends. Fizjoterapia Polska, 24(2), 56-65. DOI: https://doi.org/10.56984/8ZG5608M3Q

[7] Al-Zwainy, F., Saad, A., & Saady, A. (2024). Systematic Literature Review of The Impact of Project Management Offices In Developing The Construction Sector. Journal of Al-Azhar University Engineering Sector, 19(70), 265-273.DOI: 10.21608/auej.2024.254573.1515.

[8] Fei, W., Opoku, A., Agyekum, K., Oppon, J. A., Ahmed, V., Chen, C., & Lok, K. L. (2021). The critical role of the construction industry in achieving the sustainable development goals (SDGs): Delivering projects for the common good. Sustainability, 13(16), 9112. https://doi.org/10.3390/su13169112.

[9] Hammoody, O., Al-Somaydaii, J., Al-Zwainy, F., & Hayder, G. (2022). Forecasting and determining of cost performance index of tunnels projects using artificial neural networks. International Journal for Computational Civil and Structural Engineering, 18(1), 51-60.https://doi.org/10.22337/2587-9618-2022-18-1-51-60

[10] Hussam K. Risan, Adel A. Al-Azzawi and Faiq M.S. Al-Zwainy. (2024). Statistical Analysis for Civil Engineers: Mathematical Theory and Applied Experiment Design. (1st ed.). Elsevier Ltd, USA.

[11] Hussein, S. G., & Al-Zwainy, F. M. (2024). Diagnosing and Identifying Standards Affecting on the Ready-Mix Concrete Production Plants Performance: An Analytical Study. Tikrit Journal of Engineering Sciences, 31(1), 211-222.https://doi.org/10.25130/tjes.31.1.18

[12] Jaber, F. K., Al-Zwainy, F. M., & Hachem, S. W. (2019). Optimizing of predictive performance for construction projects utilizing support vector machine technique. Cogent Engineering, 6(1), 1685860.https://doi.org/10.1080/23311916.2019.1685860

[13] Jaber, F. K., Hachem, S. W., & Al-Zwainy, F. M. (2019). Calculating the indexes of earned value for assessment the performance of waste water treatment plant. ARPN Journal of Engineering and Applied Sciences, 14(4), 792-802.

[14] Jasim, N. A., Maruf, S. M., Aljumaily, H. S., & Al-Zwainy, F. (2020). Predicting index to complete schedule performance indicator in highway projects using artificial neural network model. Archives of Civil Engineering, 66(3).DOI: 10.24425/ace.2020.134412

[15] Kerzner, H. (2017). Project management: A systems approach to planning, scheduling, and controlling. (12th ed.). John Wiley & Sons, Inc., USA.

[16] Khan, Q., Hayder, G., & Al-Zwainy, F. M. (2023). River Water Suspended Sediment Predictive Analytics Using Artificial Neural Network and Convolutional Neural Network Approach: A Review. Sustainability Challenges and Delivering Practical Engineering Solutions: Resources, Materials, Energy, and Buildings, 51-56.https://doi.org/10.1007/978-3-031-26580-8_10

[17] Mubarak, S. A. (2015). Construction project scheduling and control. (3rd ed.). John Wiley & Sons, Inc., USA.

[18] Nabeel S. Juwaied and Faiq Mohammed Sarhan Al-Zwainy. (2017). PILE DESIGN USING MULTIPLE LINEAR REGRESSION MODEL. ARPN Journal of Engineering and Applied Sciences, 12(15), 4518-4525.

[19] Raheem, S. H., & Al-Zwainy, F. M. (2020, November). Innovation of Analytical Software for Financing Construction Projects: Infrastructure Projects in Iraq as a Case Study. In IOP Conference Series: Materials Science and Engineering (Vol. 978, No. 1, p. 012015). IOP Publishing.DOI:10.1088/1757-899X/978/1/012015

[20] Risan, H. K., Serhan, F. M., & Al-Azzawi, A. A. (2024, January). Management of a typical experiment in engineering and science. In AIP Conference Proceedings (Vol. 2864, No. 1). AIP Publishing.https://doi.org/10.1063/5.0186079

[21] Zamim, S. K., Faraj, N. S., Aidan, I. A., Al-Zwainy, F. M., AbdulQader, M. A., & Mohammed, I. A. (2019). Prediction of dust storms in construction projects using intelligent artificial neural network technology. Periodicals of Engineering and Natural Sciences, 7(4), 1659-1666.DOI: http://dx.doi.org/10.21533/pen.v7i4.857

[22] Zou, J., Kiviniemi, A., & Jones, S. W. (2017). A review of risk management through BIM and BIM-related technologies. Safety Science, 97, 88-98. https://doi.org/10.1016/j.ssci.2015.12.027