Neural Network Models for Actual Cost and Actual Duration Estimation in Construction Projects: Findings from Greece

Panagiotis Karadimos, Leonidas Anthopoulos

***Abstract*—**Predicting the actual cost and duration in construction projects concern a continuous and existing problem for the construction sector. This paper addresses this problem with modern methods and data available from past public construction projects. 39 bridge projects, constructed in Greece, with a similar type of available data were examined. Considering each project's attributes with the actual cost and the actual duration, correlation analysis is performed and the most appropriate predictive project variables are defined. Additionally, the most efficient subgroup of variables is selected with the use of the WEKA application, through its attribute selection function. The selected variables are used as input neurons for neural network models through correlation analysis. For constructing neural network models, the application FANN Tool is used. The optimum neural network model, for predicting the actual cost, produced a mean squared error with a value of 3.84886e-05 and it was based on the budgeted cost and the quantity of deck concrete. The optimum neural network model, for predicting the actual duration, produced a mean squared error with a value of 5.89463e-05 and it also was based on the budgeted cost and the amount of deck concrete.

***Keywords*—**Actual cost and duration, attribute selection, bridge projects, neural networks, predicting models, FANN TOOL, WEKA.

I. Introduction

CONSTRUCTION industry is one of the most important economic sectors that affect a country's development. It is no coincidence that a large part of a country's private and public investment is directed toward the development, maintenance and operation of facilities such as roads, bridges, buildings, wastewater treatment plants (WTPs), landfills, sewerage and electricity networks, etc. The overall contribution of the construction sector in a country's gross domestic product (GDP) growth is crucial. However, project failures in terms of meeting the planned duration and budget concern a major problem that generates economic impacts for both the project owner and the constructor. This problem is bigger in public construction projects since contract changes are harder to be performed

Construction cost and the duration estimation are being performed initially during the planning phase and are being re-estimated during the implementation phase too. Cost and time prediction is of great importance for most professionals in the construction industry and is traditionally being identified as a key success factor in a construction project [10]. However, despite the evolution of the project management science and its methods and tools, in recent decades construction projects exhibit poor performance regarding meeting the planned duration and cost [6], [7], [10]. Due to the uniqueness of each project, the prediction of a project's cost and duration has risks [20]. Serious efforts have been made and are being made to correctly predict these two factors, which have led to new methods and techniques [14].

One of the technological advancements progressively adopted by project management science is artificial intelligence (AI) and more specifically the utilization of neural networks. Applications of Artificial Neural Networks (ANN) in construction management, in general, go back to the early 1980s. These applications cover a broad range of construction issues. Neural network models have been developed internationally to assist managers and contractors in many crucial decisions. Some of these models were designed for duration and cost prediction; for decision making; predicting the percentage of the mark up and production rate etc. [14].

Bridges are crucial infrastructure projects. They belong to transportation network projects and they significantly affect their total cost and duration. The factors that affect the cost and duration of a bridge range from many quantitative (e.g., geometric characteristics, quantities of embedded materials) to several qualitative (e.g., type of bridge, construction methods used, location of construction, seismic activity).

The aim of this paper is twofold: to identify and assess the factors that correlate with the actual cost and duration of a construction project; and to produce neural network models that can estimate the final cost and duration for bridge construction projects, according to the existing data. Then, considering our sample projects, correlation results of the selected variables, as well as the impact they have on the actual cost and the actual duration of bridge projects are highlighted. Furthermore, a structured approach in producing neural network models is analyzed. The corresponding models are presented along with their performance. Finally, discussion and conclusions along with limitations and further research are being performed.

II. Project Performance Prediction Models

A lot of research work has been focused on producing models to predict the actual cost and duration of construction projects.

Aretoulis [5], after analyzing 20 highway construction

Panagiotis Karadimos is a PhD candidate and Leonidas Anthopoulos is a full Professor with the Department of Business Administration, School of Economics and Business, University of Thessaly, Geopolis Campus, GR 41500, Larissa, Greece (e-mail: pkaradimos@uth.gr, lanthopo@uth.gr).

projects in Greece, identified the critical factors affecting the actual project at completion through correlation analysis using SPSS and through WEKA application. Furthermore, several neural network models were proposed to predict the final project cost based on initial data available at the bidding stage that were lowest awarding bid, budgeted cost, technical works cost, electrical cost, tunnels, landscaping, earthmoving works cost, the number of lanes, signs, geomorphology, surfacing cost (asphalt) and paving cost (base course).

Barros et al. [9] focused on the development of an estimation technique for construction highway projects using ANN. Networks of different architectures were constructed with 10,15 and 20 neurons and they were trained and tested with the use of a backpropagation algorithm. Based on this, 14 highway projects in Brazil were considered and their data were collected and analyzed. After trials and errors, 11 parameters that contribute the most to the construction final budget were selected which were the class of the road, road extension, average transport distance of cement, average transport distance of steel, average transport distance of petroleum asphalt cement, execution time, number of bridges executed, extension of the bridges, volume of embankment, volume of excavation and volume of bituminous concrete.

Chandanshive and Kambekar [11] developed a multilayer feed-forward neural network model trained along with a backpropagation algorithm for predicting building construction cost. A dataset of 78 building construction projects was obtained from the mega-urban city of Mumbai (India) and geographically nearby regions. The most influential design parameters of the structural cost of buildings were identified and assigned as input parameters, which were ground floor area, typical floor area, number of floors, structural parking area, quantity of elevator wall, quantity of exterior wall, quantity of exterior plaster, area of flooring, number of columns, types of foundation and number of householders.

Arafa and Alqedra [4] developed an efficient model to estimate the cost of building construction projects at early stages using ANN. A database of 71 building projects was used from the construction industry of the Gaza Strip. The input layer of the ANN comprised seven parameters, namely: ground floor area, typical floor area, number of storeys, number of columns, type of footing, number of elevators, and number of rooms.

Yadav et al. [37] developed a cost estimation technique by using an ANN that could forecast the total structural cost of residential buildings by considering various parameters. In this study, data from the last 23 years were collected from the schedule of rate book (SOR) and general studies. Eight input parameters, namely, cost of cement, sand, steel, aggregates, mason skilled workers, non-skilled workers and the contractor per square feet construction were selected.

Roxas and Ongpeng developed an ANN model that can predict the total structural cost of building projects in the Philippines [31]. Data from 30 building projects were collected. Six input parameters were used, namely: number of storeys, number of basements, floor area, volume of concrete, area of formworks, and weight of reinforcing steel. The feedforward backpropagation technique was used to generate the best model for the total structural cost.

Considering again the cost concept, Cheng et al. [12] proposed an evolutionary fuzzy hybrid neural network (EFHNN) for projects in the construction industry. The research focus was on conceptual cost estimates. The same research team, Cheng et al. [13] developed an EFHNN to enhance project cash flow management.

Jiang and Wu [23], after analyzing 1,818 projects from Indiana's area (Indiana Department of Transportation), reported that factors that affect directly the duration and the cost of each project include: type of the project, project size, manpower, equipment, construction management, location of the project and weather conditions.

Wang et al. [35] developed ANN ensemble and support vector machine classification models to predict the project cost and schedule success. The model input data were based on the status of early planning. Their survey managed to record 92 building projects.

Stoy et al. [33] focusing on German residential building projects produced a neural network predictive model for the most relevant cost element, namely: external walls. Multiple linear regression methodology is employed for the final parameter estimation.

Aziz [8] developed a tool based on statistical regression analysis, which except from predicting the cost and time of a project would, also, help to evaluate the project's performance during construction. In this study, three methods of analysis were used: ridge regression analysis, general regression analysis and nonlinear partial least-square regression analysis. Data collected from completed projects included the type of pavement, contract value, duration and project miles.

Kang and Kim [25] predict the risk cost and the bidding price of a plant construction project based on the surveyed risk information. The model structure was designed and a prototype program was developed for the analysis. The proposed model and the prototype program succeeded in properly estimating the bidding price, considering the risk cost in plant construction projects.

Sonmez [32] applied neural networks with bootstrap prediction intervals for range estimation of construction costs. Here, neural networks are used for modelling the mapping function between the factors and costs. The bootstrap method is used to quantify the level of variability included in the estimated costs. This methodology is implemented to range estimation of building projects. The proposed integrated approach accounts for an alternative for conceptual estimation of costs.

Koo et al. [26] focused on Case-Based Reasoning (CBR) approaches. The aim of their research was based on the development of a construction cost prediction model using the advanced CBR approach. Their sample included 101 cases of multi-family housing projects. The proposed approach managed to integrate successfully CBR, MRA (Multiple Regression Analysis), ANN, and the optimization process using a genetic algorithm. The approach succeeded in supporting the stakeholders in charge of predicting and managing a construction cost in the early stages of a construction project to obtain more accurate results from historical cases.

Titirla and Aretoulis [34], after analyzing 37 highway projects in Greece identified the factors affecting the actual project duration at completion through correlation analysis using SPSS and through the WEKA application. Furthermore, several network models were proposed to predict the project construction duration based on initial data available at the bidding stage such as lanes, initial duration, length, initial cost, tunnels, technical projects, bridges, land requirement, embankment, landfill, geotechnical projects and tender offer [34].

Using 168 building projects constructed in Spain, the MRA was used by Guerrero et al. [19] for developing a prediction model that allows the estimation of project duration of new builds. The proposed model uses as predictor variables the following variables: the number of floors, project type, gross floor area (GFA) and the cost/GFA relationship. In this research, the logarithmic form of construction speed was identified as the most appropriate response variable. Both GFA and cost are necessary to achieve a prediction model with the highest accuracy; however, GFA has greater influence than the cost on project duration.

The aim of the Al-Saidi et al. [1] was to predict the construction duration of road projects in the Republic of Iraq. Historical data were adopted for 99 projects for the interval between 2000 to 2017 from the Roads and Bridges Directorate (RBD). ANN model was used to estimate the duration using six variables (length of road, No.of lane, No.of intersection, volume of earth, type of pavement and furniture level).

Glymis et al. [18] proposed three, selected, neural network models for predicting actual project duration for highways, based on tender budget, length of the highway project, number of lanes, number of technical projects, number of bridges, tunnels and road total length.

Petruseva et al. [30] produced a neural network model for predicting construction project duration. Key data of the total of 75 buildings constructed in the Federation of Bosnia and Herzegovina have been collected through field studies. The collected data contained information for the contracted and real-time construction, there were also data for the use of these 75 projects and the construction year. For predicting the construction time, a linear regression model using “time-cost” was applied to these data. Then, to the same data, a multilayer perceptron neural network (MLP-NN) predictive model was applied and a significant improvement of the accuracy of the forecasting was obtained.

Marzoughi et al. [28] focused their research on introducing a decision support framework for estimating project duration under the influence of the weather. Their study proposed a five-module framework that integrated weather variables, project performance variables and duration of project activities Their approach implemented expert knowledge relating to the importance of variables of weather, pairwise comparisons of weather variables with respect to different criteria of performance and pairwise comparisons of performance variables concerning project activities. The study moved on to produce a model using multivariate statistical techniques and an analytical network process (ANP) to estimate the duration of project activities considering the impact of weather.

Gab-Allah et al. [17] developed an ANN model for predicting the expected construction duration of building projects during their preliminary studies, where no detailed planning is available. The program used was MATLAB and it was used as a suitable environment for developing the proposed model. The required data were collected from 130 building projects in Egypt, which fall within the appropriate sample size.

Based on variables that were known at the planning stage, such as contract type, project type and planned cost, Irfan et al. [22] investigated the prediction of the project duration of highways. This study focused on the mathematical relationships between the project duration of highways and the contract type as well as the project type and the magnitude of the planned cost. The main findings suggested that all other factors remaining the same, the duration of fixed-date deadline contracts generally exceeded that of fixed-duration contracts. Furthermore, the paper highlighted that higher levels of planned cost translated non-linearly into greater project duration.

The application of the ANN to predict the duration of implementation of a residential construction project from the pre-design stage to completion is comprehensively discussed in [2]. The study applied the back-propagation (BP) network made of nodes for error evaluation of the training states.

Liu [27] divided the parameters in highway construction into two groups: weather and other factors. According to the different characteristics of these two groups of random factors, the study introduced different methods to deal with them when estimating the work duration. The computer simulation technique was used to estimate the effect of weather conditions. PERT method was implemented to estimate the effect of the other random factors.

Antoine et al. [3] focused on investigating the relationship between project duration, project intensity and timing of cost certainty in highway project delivery methods. They concluded that alternative contracting methods are viable options for shortening project durations, establishing early cost certainty during project delivery and delivering projects at a more intense pace. More specifically, alternative contracting methods of construction manager/general contractor CM/GC and Design-Build are superior to the traditional Design Bid Build method for the project performance.

Migliaccio and Shrestha [29] analyzed design-build (DB) procurement activities’ durations for highway projects. Results revealed that project size measured by contract dollar amount affects the duration of DB procurement activities. Procurement durations did not correlate with the project cost for projects costing less than \$250m. On the other hand, when the project cost is higher than \$250m, a linear correlation between these two variables appeared.

Hosseinian and Reinschmidt [21] aimed at finding the best progress model for tunneling projects with the new Austrian tunneling method (NATM) by conducting Bayesian analysis on available data of a massive project. The analysis revealed that the dual Gompertz function was the most reliable model for this purpose. The results of this research bring advantages to future NATM tunnel constructions.

III. Research Methodology

The main aim of the current research is to produce reliable and efficient neural network models for the prediction of actual cost and actual duration, regarding the construction of bridges, with emphasis on the special characteristics of Greece. The sample projects under examination include 39 bridge projects. For these specific projects, it became possible to record a detailed amount of the same type of data, both quantitative and qualitative. The initial relevant research approaches were mainly based on multiple linear regression methods. The limitations of these methods focused on their linear and parametric nature [5]. Furthermore, available data also included qualitative parameters and the neural networks seemed to be a better approach than a 'traditional' multiple linear regression analysis [5]. Moreover, [24] emphasizes that neural networks possess the features of the ability to learn and generalize the acquired knowledge, the ability to adapt to changing conditions and small sensitivity to errors in the input data. The latter is crucial, as there is always an issue regarding the data reliability, especially when it comes to cost and duration. Based on these facts and the international literature, the current paper implements ANN to predict the actual cost and the actual duration for bridge projects constructed in Greece.

A. Steps of Methodological Approach

The proposed methodology is based on three tools, namely: FANN tool (Fast Artificial Neural Network Tool) for neural network implementation, WEKA (Waikato Environment for Knowledge Analysis) for attribute selection and SPSS (Statistical Package for the Social Sciences) for correlation analysis. More specifically, the FANN tool was used to facilitate the implementation of an abundance of different neural network libraries. SPSS along with WEKA facilitates variables' screening [5]. In essence, the FANN tool was used to produce a large amount of different ANN models. SPSS was used for data description and running correlation analysis to identify the predictive capability of each independent variable and rank them based on their correlation coefficients. Then, WEKA application was used to identify a subgroup of variables both within the initially available variables but also within the group of highly correlated variables as identified by the correlation analysis. The methodology steps included the following [5]:

Step1. Consideration of 39 bridge projects and collection of the corresponding data.

Step2. The construction of an appropriate SPSS database, including both quantitative and qualitative variables.

Step3. Descriptive statistics of the sample projects' variables.

Step4. Correlation statistical analysis among the available variables and actual project cost and also among the available variables and actual project duration. Analysis was conducted one time for quantitative variables and the second time for variables, both quantitative and qualitative in type.

Step5. Creation of ranked lists of variables for dependent variables of actual cost and actual duration, respectively, based on decreasing degree of correlation for potential input neurons.

Step6. Proposal of neural network models for dependent variables of actual cost and actual duration, respectively, based on FANN Tool. The models were created starting with the highest correlating variable and then adding one more variable each time from the ordered list, based on the degree of correlation (step 5).

Step7. Additional screening of the correlated variables with WEKA application and identification of the most efficient sub-group of variables for neural network inputs.

Step8. Proposal of neural network models, based on FANN Tool, for predicting the actual cost and actual duration of bridge construction projects based on the variables identified in step 7.

The methodological approach for the actual cost prediction is graphically depicted in Fig. 1 [5]. The same applies for predicting the actual duration.

B. Considered Variables

An SPSS database was organized to record all the available variables. The variables are characterized as quantitative and qualitative. The quantitative variables take on numerical values, and the qualitative variables take on binary (Yes/No) or ordinal values. Table I depicts the available variables: The explanation and type of each considered variable are as follow:

- Single/Twin bridge: This is a quantitative variable and concerns whether a bridge is single or twin. The values it takes are '1' for single bridges and '2' for twin bridges.
- Deck length: This is a quantitative variable and refers to the total length of a bridge's deck in meters (m).
- Bridge surface: This is a quantitative variable and concerns the total area of a bridge's surface in square meters (m^2).
- Maximum height of piers: This is a quantitative variable, it concerns the maximum height of piers, which in most cases concerns the height of the piers positioned in the middle of the deck, in meters (m).
- Quantity of deck concrete: This is a quantitative variable and concerns the amount of concrete required for constructing a bridge's deck, in cubic meters (m^3).
- Budgeted Cost: This is a quantitative variable and concerns the amount of the estimated construction cost of a bridge project at the planning stage in euros (€). To make the economic data of bridges comparable, the budget of each bridge was adjusted to values of the 1/1/2020 with the application of the corresponding inflation rates on the initial project budget estimation and

gradually for each year, from the year of completion of the planning stage until the end of 2019. The inflation rates were based on the official data of the Hellenic Statistical Office (EL.STAT) [20].

- Actual duration of construction: This is a quantitative variable and concerns the duration required for constructing a bridge, from the start of the works until their completion, in weeks.
- Actual cost of construction: This is a quantitative variable and concerns the actual cost required for constructing a bridge from the start of the works until their completion in Euros (€). To make the economic data of bridges comparable, the actual cost of each bridge was adjusted to values of the 1/1/2020 with the application of the corresponding inflation rates and gradually for each year, from the year of completion of the project until the end of 2019. The inflation rates were based on the official data of the Hellenic Statistical Office (EL.STAT) [20].
- Region: This is a qualitative variable that takes values from 1 to 7 and refers to the region in which a bridge was built (Eastern Macedonia & Thrace = 1, Central Macedonia = 2, Western Macedonia = 3, Epirus = 4, Thessaly = 5, Central Greece = 6, Peloponnese = 7).
- Type of bridge deck: This is a qualitative variable and refers to the classification of a bridge regarding the type of its deck. It takes values from 1 to 3 (Deck slab = 1, Beam cross-section = 2, Box-girder cross-section = 3).
- Bridge deck construction method: This is a qualitative variable and concerns the classification of a bridge concerning the method used for constructing its deck. It takes values from 1 to 5. (Balanced Cantilever = 1, Incremental Launching = 2, Traditional scaffolding = 3, Precast beams = 4, Travelling formwork = 5).
- Piers' construction method: This is a qualitative variable and concerns the classification of a bridge regarding the method was used for the construction of its piers. It takes values from 1 to 2 (Climbing formwork method = 1, Slip forming method = 2).
- Seismic hazard zone: This is a qualitative variable and concerns the seismic hazard zone of the site where a bridge was constructed. It takes values from 1 to 3 (I = 1, II = 2, III = 3).

C. Sample and Database Description

The database includes data from 39 bridges constructed in mainland Greece. 29 of them concern highway bridges and ten rail bridges. 19 of them concern bridges of Egnatia Motorway, seven concern bridges of Kentriki Odos – E65 Motorway, ten concern bridges of Greek Railways (OSE S.A.), two concern bridges of the National Road (Ethniki Odos) and one from Olympia Odos – A8 Motorway. The main structural material of their construction was concrete (reinforced-prestressed). The bridges were selected mainly based on the availability and uniformity of their data for the complete database creation possible. For these projects, it was possible to record and collect quantitative and qualitative data of the same type. All the considered bridges were constructed between 1995 and 2015. The number of projects per region is as follows:

- Eastern Macedonia and Thrace: 2 projects
- Central Macedonia: 5 projects
- Western Macedonia: 10 projects
- Epirus: 7 projects
- Thessaly: 7 projects
- Central Greece: 4 projects
- Peloponnese: 4 projects

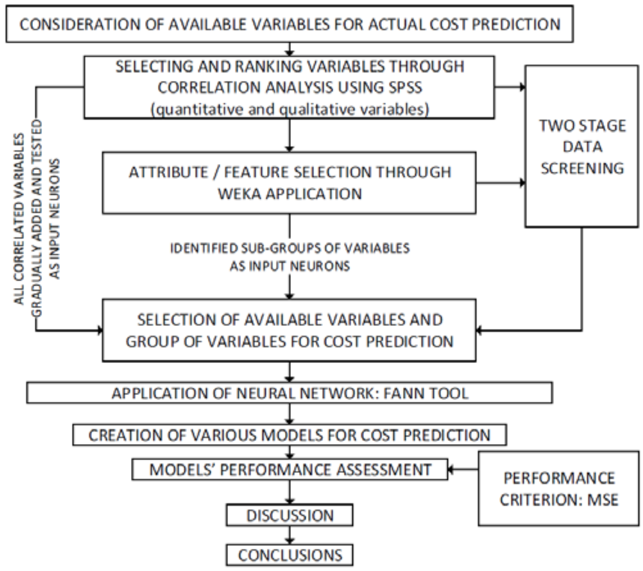

Fig. 1 Methodological approach [5]

Then, a database was created using SPSS to proceed in the subsequent analyses. The database consisted of 39 cases (number of projects) and 13 variables that cover commonly available data across all projects. These variables were recorded as quantitative and qualitative. Descriptive statistics of the sample are included in Tables II A and B.

D. Correlation Analysis for Quantitative Variables

For the determination of the correlation between the quantitative independent variables and the dependent variables of the 'Actual Cost' and 'Actual Duration' of a bridge project, the IBM SPSS statistical package was used with the bridge data table adjusted to the requirements of the application. Subsequently, and based on the findings of the correlation analysis, several neural networks were constructed in order to develop models to predict the actual cost of a bridge project and the actual duration of construction. According to Field [16], the Pearson correlation coefficient and the significance factor are the indicating factors for the assessment of the correlation analysis results. The Pearson coefficient takes on values close to 1 for strong relationships and to -1 for adverse strong relationships. Also, significance values less than 0.05 reveal a strong correlation and those values that range among 0.05 and 0.06 demonstrate the tendency to correlate. When the significance factor is less than 0.01 ($p < 0.01$), it is denoted by **, while when it is between 0.01 and 0.05 ($0.01 < p < 0.05$), it is denoted by *. Tables III and IV depict the correlation analysis results for the dependent variables of the 'Actual

Cost' and 'Actual Duration' respectively, in descending order.

Table III indicates that all independent variables have a high degree of correlation with the dependent variable of the 'Actual Cost,' with the exception of the 'Single/Twin bridge' variable, which has a weaker correlation with the dependent variable based on Pearson coefficient and significance factor. Table IV shows the high degree of correlation of all independent quantitative variables with the dependent of the 'Actual Duration'.

TABLE I
VARIABLES INCLUDED IN THE SPSS DATABASE

Quantitative variables	Qualitative variables
Single/Twin bridge	Region
Deck length (m)	Type of bridge deck
Bridge surface (m^2)	Bridge deck construction method
Maximum height of piers (m)	Piers' construction method
Quantity of deck concrete (m^3)	Seismic hazard zone
Budgeted cost (€)	
Actual duration (weeks)	
Actual cost (€)	

TABLE II A
SAMPLE DESCRIPTIVE STATISTICS - A

Variables	N	Range	Minimum
Single/Twin bridge	39	1,00	1.00
Deck length (m)	39	1,775.90	64.10
Bridge surface (m^2)	39	25,190.50	56.50
Maximum height of piers (m)	39	82.00	6.00
Quantity of deck concrete (m^3)	39	21,631.15	522.45
Budgeted cost (€)	39	55,736,099.15	538,718.78
Actual duration (weeks)	39	134.00	32.00
Actual cost (€)	39	42,457,815.92	649,311.36
Region	39	6.00	1.00
Type of bridge deck	39	2.00	1.00
Bridge deck construction method	39	4.00	1.00
Piers' construction method	39	1.00	1.00
Seismic hazard zone	39	1.00	1.00

TABLE II B
SAMPLE DESCRIPTIVE STATISTICS - B

Variables	Maximum	Mean	Std. Deviation
Single/Twin bridge	2.00	1.49	0.51
Deck length (m)	1,840.00	366.55	356.36
Bridge surface (m^2)	25,760.00	4,886.22	5,020.52
Maximum height of piers (m)	88.00	22.67	17.73
Quantity of deck concrete (m^3)	22,153.00	4,349.09	4,311.37
Budgeted cost (€)	56,274,817.93	8,224,009.31	10,369,657.18
Actual duration (weeks)	166.00	76.62	38.17
Actual cost (€)	43,107,127.28	6,982,717.14	8,383,744.93
Region	7.00	4.03	1.68
Type of bridge deck	3.00	2.28	0.89
Bridge deck construction method	5.00	2.51	1.32
Piers' construction method	2.00	1.28	0.46
Seismic hazard zone	2.00	1.46	0.51

TABLE III
CORRELATION ANALYSIS OF QUANTITATIVE VARIABLES WITH DEPENDENT VARIABLE THE 'ACTUAL COST'

A/A	Quantitative Variables	Pearson Coefficient
1	Budgeted cost	0.989**
2	Quantity of deck concrete	0.933**
3	Bridge surface	0.933**
4	Deck length	0.918**
5	Actual duration	0.748**
6	Maximum height of piers	0.595**
7	Single/Twin bridge	0.339*

TABLE IV
CORRELATION ANALYSIS OF QUANTITATIVE VARIABLES WITH DEPENDENT VARIABLE THE 'ACTUAL DURATION'

A/A	Quantitative Variables	Pearson Coefficient
1	Deck length	0.878**
2	Bridge surface	0.872**
3	Quantity of deck concrete	0.868**
4	Actual cost	0.748**
5	Budgeted cost	0.731**
6	Single/Twin bridge	0.576**
7	Maximum height of piers	0.558**

E. Correlation Analysis for both Quantitative and Qualitative Variables

Here, the qualitative variables are also included, and the correlation analysis results are presented in Tables V and VI for dependent variables of the 'Actual Cost' and 'Actual Duration' respectively, in descending order.

Table V shows the high degree of correlation for the most independent variables with respect to the dependent of the 'Actual Cost'. The high correlated independent variables are the following: Budgeted cost, quantity of deck concrete, bridge surface, deck length, actual duration, piers' construction method and maximum height of piers.

Table VI shows the high degree of correlation for the most independent variables with respect to the dependent of the 'Actual Duration'. The high correlated independent variables are the following: Deck length, bridge surface, quantity of deck concrete, actual cost, budgeted cost, single/twin bridge and maximum height of piers.

TABLE V
CORRELATION ANALYSIS OF QUANTITATIVE AND QUALITATIVE VARIABLES WITH DEPENDENT VARIABLE THE 'ACTUAL COST'

A/A	Quantitative + Qualitative Variables	Pearson Coefficient
1	Budgeted cost	0.989**
2	Quantity of deck concrete	0.933**
3	Bridge surface	0.933**
4	Deck length	0.918**
5	Actual duration	0.748**
6	Piers' construction method	0.627**
7	Maximum height of piers	0.595**
8	Type of bridge deck	0.436**
9	Single/Twin bridge	0.339*
10	Bridge deck construction method	-0.107
11	Seismic hazard zone	-0.159
12	Region	-0.278

TABLE VI
CORRELATION ANALYSIS OF QUANTITATIVE AND QUALITATIVE VARIABLES WITH DEPENDENT VARIABLE THE 'ACTUAL DURATION'

A/A	Quantitative + Qualitative Variables	Pearson Coefficient
1	Deck length	0.878**
2	Bridge surface	0.872**
3	Quantity of deck concrete	0.868**
4	Actual cost	0.748**
5	Budgeted cost	0.731**
6	Single/Twin bridge	0.576**
7	Maximum height of piers	0.558**
8	Piers' construction method	0.431**
9	Type of bridge deck	0.390*
10	Bridge deck construction method	0.060
11	Seismic hazard zone	-0.209
12	Region	-0.308

IV. APPLICATION OF NEURAL NETWORKS

For the production of the neural networks, the FANN Tool application was implemented, which is free-to-use software. One of the following learning algorithms can be selected by the user: FANN Train Incremental (gradually increasing), FANN Train Batch (clustering), FANN Train Rprop (Resilient backpropagation) and FANN Train Quickprop [18]. The basic functions of the software are the following.

- Neural Network = > Detect = > Optimum training algorithm: The possible training algorithms are used for certain epochs. The weight initialization is identical and the other parameters are fixed. The algorithm with the lowest MSE is selected [15].
- Neural Network = > Detect = > Optimum activation functions: The possible activation functions are used for certain epochs. The weight initialization is identical and the other parameters are fixed. The activation function with the lowest MSE is selected [15].
- Neural Network = > Train = > Normal: Fixed topology training. The topology and the size of the neural network are determined in advance and the weights are altered by the training to minimize the difference between the actual output values and the desired output values [15].
- Neural Network = > Train = > Cascade: Evolving topology training. The training starts with an empty ANN that consists only of output and input neurons. Hidden connections and neurons are added during training to reach the same goal as for fixed topology training [15].

Neural networks produced several models. The dependent variables were in the first case the 'Actual cost' in Euros and the second case the 'Actual duration' in weeks. 23 projects were used for training the neural networks and 16 for testing the produced models. The methodology involving the application of neural networks for predicting the actual cost and duration is based on the design and application of multiple neural networks specifically for the problem at hand.

The current research is also contributing and focusing on identifying the best possible combination of input variables for the optimum prediction result. Regarding the structure of the neural network and the relevant equations, these are left to the application itself to search and identify in order to define the optimum design and used parameters, each time, through its 'cascade' function.

A. Neural Network Models Based on Quantitative Variables

The pool of available variables for neural network construction for predicting the 'Actual Cost' is depicted in Table III, while for predicting the 'Actual Duration' is depicted in Table IV. The neural network models are presented in Tables VII and VIII for dependent variables of the 'Actual Cost' and 'Actual Duration' respectively, along with the input variables and the mean squared error (MSE). The models are ranked in increasing order of MSE value. The model with the lowest MSE is the most effective. The top correlated variable was used as a single input neuron in the first model that was applied in both cases. Then, each consecutive model was realized by adding every time a new, additional, variable. The addition of new variables was following the correlation coefficient-based ranking of variables. The resulting models can be seen in Tables VII and VIII for predicting the 'Actual Cost' and 'Actual Duration' respectively.

In the case where the dependent variable was the 'Actual Cost', 13 models were produced (Model 1 to Model 13). The best performing model is Model No 8 which includes the top two most correlated variables, namely: Budgeted cost, Quantity of deck concrete. In the case where the dependent variable was the 'Actual Duration', 13 models were produced (Model 18 to Model 30). The best performing model is Model No 25 which includes the top two most correlated variables, namely: Deck length, bridge surface.

B. Neural Network Models Based on Quantitative and Qualitative Variables

The pool of available variables for neural network construction for predicting the 'Actual Cost' is depicted in Table V, while for predicting the 'Actual Duration' is depicted in Table VI. The neural network models are presented in Tables IX and X for dependent variables of 'Actual Cost' and 'Actual Duration' respectively, along with the input variables and the MSE. The models are ranked in increasing order of MSE value.

The model with the lowest MSE is the most effective. The top correlated variable was used as a single input neuron in the first model that was applied in both cases. Then, each consecutive model was realized by adding every time a new, additional, variable. The addition of new variables was following the correlation coefficient-based ranking of variables. The resulting models can be seen in Tables IX and X for the prediction of 'Actual Cost' and 'Actual Duration' respectively.

In the case where the dependent variable was the 'Actual Cost', 17 models were produced (Model 1 to Model 17). Models 1-14 are the same as the ones in the case where only quantitative input variables were considered. The best performing model is Model No 8 which includes the top two most correlated variables, namely: Budgeted cost, the quantity of deck concrete. In the case where the dependent variable was the ‘Actual Duration’, 14 models were produced (Model 18 to Model 31). The best performing model is Model No 31 which includes the top eight most correlated variables, namely: Deck length, bridge surface, the quantity of deck concrete, actual cost, budgeted cost, single/twin bridge, the maximum height of piers, piers’ construction method.

TABLE VII
NEURAL NETWORK MODELS BASED ON QUANTITATIVE VARIABLES WITH DEPENDENT VARIABLE THE ‘ACTUAL COST’

Model	Combination	Input Variables	MSE
8	Top two	Budgeted cost + Quantity of deck concrete	3.84886e-05
9	Top three	Budgeted cost + Quantity of deck concrete + Bridge surface	3.8775e-05
10	Top four	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length	4.48561e-05
2	Single second	Quantity of deck concrete	5.15374e-05
3	Single third	Bridge surface	7.98023e-05
5	Single fifth	Actual duration	8.3318e-05
12	Top six	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Maximum height of piers	8.53354e-05
11	Top five	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration	9.01897e-05
4	Single fourth	Deck length	9.02971e-05
13	Top seven	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Maximum height of piers + Single/Twin bridge	9.08738e-05
1	Single top	Budgeted cost	9.42549e-05
6	Single sixth	Maximum height of piers	0.00669171
7	Single seventh	Single/Twin bridge	0.00947128

TABLE VIII
NEURAL NETWORK MODELS BASED ON QUANTITATIVE VARIABLES WITH DEPENDENT VARIABLE THE ‘ACTUAL DURATION’

Model	Combination	Input Variables	MSE
25	Top two	Deck length + Bridge surface	8.19725e-05
27	Top four	Deck length + Bridge surface + Quantity of deck concrete + Actual cost	8.654986e-05
30	Top seven	Deck length + Bridge surface + Quantity of deck concrete + Actual cost + Budgeted cost + Single/Twin bridge + Maximum height of piers	8.69428e-05
26	Top three	Deck length + Bridge surface + Quantity of deck concrete	9.26152e-05
28	Top five	Deck length + Bridge surface + Quantity of deck concrete + Actual cost + Budgeted cost	9.76806e-05
20	Single third	Quantity of deck concrete	9.88862e-05
18	Single top	Deck length	9.94046e-05
29	Top six	Deck length + Bridge surface + Quantity of deck concrete + Actual cost + Budgeted cost + Single/Twin bridge	9.98832e-05
21	Single fourth	Actual cost	0.000140701
22	Single fifth	Budgeted cost	0.000511923
19	Single second	Bridge surface	0.00084067
24	Single seventh	Maximum height of piers	0.00405333
23	Single sixth	Single/Twin bridge	0.014457

TABLE IX
NEURAL NETWORK MODELS BASED ON QUANTITATIVE AND QUALITATIVE VARIABLES WITH DEPENDENT VARIABLE THE ‘ACTUAL COST’

Model	Combination	Input Variables	MSE
8	Top two	Budgeted cost + Quantity of deck concrete	3.8488e-05
9	Top three	Budgeted cost + Quantity of deck concrete + Bridge surface	3.8775e-05
10	Top four	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length	4.4856e-05
2	Single second	Quantity of deck concrete	5.1537e-05
16	Top eight	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Piers’ construction method+ Maximum height of piers + Type of bridge deck	7.4098e-05
15	Top seven	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Piers’ construction method + Maximum height of piers	7.5588e-05
3	Single third	Bridge surface	7.9802e-05
5	Single fifth	Actual duration	8.3318e-05
12	Top 5 + seventh	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Maximum height of piers	8.5335e-05
11	Top five	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration	9.0189e-05
4	Single fourth	Deck length	9.0297e-05
13	Top five + seventh + eighth	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Maximum height of piers + Single/Twin bridge	9.0873e-05
1	Single top	Budgeted cost	9.4254e-05
14	Top six	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Piers’ construction method	9.6269e-05
17	All variables	Budgeted cost + Quantity of deck concrete + Bridge surface + Deck length + Actual duration + Piers’ construction method + Maximum height of piers + Type of bridge deck + Single/Twin bridge	9.7435e-05
6	Single seventh	Maximum height of piers	0.00669171
7	Single nineth	Single/Twin bridge	0.00947128

V. WEKA APPLICATION FOR ATTRIBUTE SELECTION

WEKA is described as a collection of machine learning algorithms that are used for data mining tasks. WEKA contains tools for data pre-processing, clustering, regression, visualization and association rules. It is also applicable for the development of new machine learning schemes [36]. In essence, WEKA is a machine learning software in Java. A research team in Waikato University, in New Zealand, has incorporated several standard machine learning (ML) techniques into a software “workbench” called WEKA. The name WEKA stands for Waikato Environment for Knowledge Analysis. It is an extremely efficient tool that enables a specialist in a particular field to use ML to derive useful knowledge from databases that are far too large to be analyzed by hand. The users of WEKA include ML researchers and industrial scientists. Teaching is also a field where WEKA has been widely acknowledged [36].

TABLE X
NEURAL NETWORK MODELS BASED ON QUANTITATIVE AND QUALITATIVE VARIABLES WITH DEPENDENT VARIABLE THE 'ACTUAL DURATION'

Model	Combination	Input Variables	MSE
31	Top eight	Deck length + Bridge surface + Quantity of deck concrete + Actual cost + Budgeted cost + Single/Twin bridge + Maximum height of piers + Piers' construction method	7.92923e-05
25	Top two	Deck length + Bridge surface	8.19725e-05
27	Top four	Deck length + Bridge surface + Quantity of deck concrete + Actual cost	8.65498e-05
30	Top seven	Deck length + Bridge surface + Quantity of deck concrete + Actual cost + Budgeted cost + Single/Twin bridge + Maximum height of piers	8.69428e-05
26	Top three	Deck length + Bridge surface + Quantity of deck concrete	9.26152e-05
28	Top five	Deck length + Bridge surface + Quantity of deck concrete + Actual cost + Budgeted cost	9.76806e-05
20	Single third	Quantity of deck concrete	9.88862e-05
18	Single top	Deck length	9,94046e-05
29	Top six	Deck length + Bridge surface + Quantity of deck concrete + Actual cost + Budgeted cost + Single/Twin bridge	9.98832e-05
21	Single fourth	Actual cost	0.000140701
22	Single fifth	Budgeted cost	0.000511923
19	Single second	Bridge surface	0.00084067
24	Single seventh	Maximum height of piers	0.00405333
23	Single sixth	Single/Twin bridge	0.014457

A. Quantitative Variables

The WEKA application was used to identify subgroups of critical variables. The chosen evaluator was: 'CfsSubsetEval-P1-E1' and the search method: 'BestFirst-D1-N5'. For the case where the 'Actual Cost' was taken as the dependent variable, the attributes that were considered, seven in number, included: Budgeted cost, quantity of deck concrete, bridge surface, deck length, actual duration, the maximum height of piers and single/twin bridge. WEKA identified a sub-group of five attributes that included: Deck length, bridge surface, the maximum height of piers, quantity of deck concrete and single/twin bridge. These five selected attributes were used as input neurons for creating Model No 32. Application of this neural network returns an MSE equal to 7.264476e-005. This is the fifth best model that includes quantitative variables.

For the case where the 'Actual Duration' was taken as the dependent variable, the attributes that were considered, seven in number, included: Budgeted cost, quantity of deck concrete, bridge surface, deck length, actual cost, the maximum height of piers and single/twin bridge. WEKA identified a sub-group of two attributes that included: Budgeted cost and the quantity of deck concrete. These two selected attributes were used as input neurons for creating Model No 33. Application of this neural network returns an MSE equal to 5.89463e-005. This is the best model that includes quantitative variables. The combination of the two attributes that give the best model (No 33) with dependent variable the 'Actual Duration' is the same as the combination of the same attributes that give the best model with dependent variable the 'Actual Cost'.

The WEKA application was not used for independent quantitative + qualitative variables for both dependent variables due to the large number of possible values that each independent variable could get. That could lead to a large number of qualitative variables which could increase the risk of unreliable results by the WEKA process.

VI. DISCUSSION - CONCLUSIONS

In this paper, with the methodological approach followed, a significant number of effective and reliable models of neural networks were constructed regarding the prediction of the actual cost and also the actual duration of bridge projects in Greece. The results given by the models produced are very interesting and they could be compared with other prediction methods that take different approaches to predict bridge project quantities.

Part of the methodological approach followed concerned the correlation analysis of the variables with the use of the IBM SPSS application. The aim was to determine the degree of correlation between the independent variables (quantitative and qualitative) with the dependent variables, which in one case was the actual cost of a bridge and in the second case was the actual duration of a bridge. Regarding the results of the correlation analysis, the obtained values of the Pearson Coefficient and their comparisons were of great interest. Concerning the correlation of the independent variables with the dependent variable of the actual cost, what has emerged is that the most correlated independent variable was the budgeted cost followed by the quantity of deck concrete and the bridge surface. This applies to both cases considered, of quantitative independent variables and the added quantitative and qualitative independent variables too. In the case where the dependent variable was the actual duration of construction, what has emerged is that the most correlated independent variable was the deck length followed by the bridge surface and the quantity of deck concrete. This applies to both cases considered, of quantitative independent variables and the added quantitative and qualitative variables too.

According to the methodological approach followed, the next step involved the production of neural network models. Initially, the models were based on quantitative variables. In the case where the dependent variable was the actual cost, the best-performing model was the one that included the two most correlated variables which were the budgeted cost and the quantity of deck concrete. The next most efficient model was the one that included the three most correlated variables, which, in addition to the previous two, the bridge surface was added. The third most efficient model was the one that included the four most correlated variables. In other words, in addition to the previous three variables of the second most efficient model, the deck length was added. In case that the actual duration was the dependent variable, the best-performing model was the one that included the two most correlated variables, the deck length and the bridge surface. The next most efficient model was the one that included the four most correlated variables, which were, in addition to the previous two, the quantity of deck concrete and the actual cost. The third most efficient model was the one that included all seven quantitative variables which were evaluated for their correlation with the dependent variable. These are, in addition to the previous four of the second most efficient model, the budgeted cost, single/twin bridge and the maximum height of piers.

Regarding the models which were based on the quantitative and qualitative variables and in the case the dependent variable was the actual cost, the three most efficient models were the same as the models that emerged from the evaluation of the quantitative variables. In the case where the dependent variable was the actual duration, the most efficient model was the one that included the following variables: Deck length, bridge surface, the quantity of deck concrete, actual cost, budgeted cost, single/twin bridge, the maximum height of piers and the piers' construction method. In second and third place were the same models which were, respectively, in first and second place in the evaluation of quantitative variables.

The WEKA application was used only for quantitative variables. In the case where the actual cost was considered as the dependent variable, the result was the selection of a subgroup of five variables which were: Single/twin bridge, deck length, bridge surface, the maximum height of piers and the quantity of deck concrete. These variables were used as input neurons and the model constructed was the fifth most efficient among those of quantitative variables as well as among those of quantitative and qualitative variables. In the case where the actual duration of construction was considered as the dependent variable, the result was the selection of a subgroup of two variables which were the following: the quantity of deck concrete and the budgeted cost of the project. These variables were used as input neurons and the model constructed was the most efficient among those of the quantitative variables as well as those of the quantitative and qualitative variables. The WEKA application was not used for independent quantitative + qualitative variables for both dependent variables due to the large number of possible values that each independent variable could get. That could lead to a large number of qualitative variables which could increase the risk of unreliable results by the WEKA process.

The models created, in their majority, are regarded as reliable and of high performance concerning the prediction of the bridge quantities they were considered, since in most of them the MSE was very small, of the order of 10^{-5}. More specifically, considering all the models (quantitative and qualitative + WEKA combinations) which were created taking as dependent variable the actual cost, 16 out of 18 models were of the order of 10^{-5}. Considering all the models (quantitative and qualitative + WEKA combinations) which were created taking as dependent variable the actual duration, ten out of 15 models were of the order of 10^{-5}.

The present study identified, through correlation analysis, using the statistical software package IBM SPSS as well as the WEKA application, the critical quantities and combinations of them that affect the actual cost and the actual duration of bridge projects. Besides, a significant number of neural network models were proposed which predict the actual cost and the actual duration of bridge projects based on a database created from previous projects. The models created and the results of the research which emerged could be useful to bridge construction contractors, design firms, consulting firms, local government agencies as well as technical tendering authorities of construction projects. Also, the models created and the results of the research can be useful and workable during the planning stage of a bridge, at the tender stage, as well as during construction by facilitating the control and management of project funding as well as the control and management of time schedules.

It should be emphasized that other parameters can also affect the actual cost and the actual duration of a bridge which are difficult to predict, quantify and generally assess their share in the construction of a bridge. Many times, many participants in the construction of a bridge ignore them with unforeseen consequences for the successful construction of a project. Factors such as archaeological findings, project financing, weather conditions, environmental permits and expropriations are some of those that can cause an increase in the actual cost as well as the actual duration of a bridge.

In our time, technological developments combined with the high experience gained in modern and complex projects, allow the successful planning and design of a project from the early stages. With the approach of the neural networks followed in this research, the actual cost and the actual duration can be easily estimated using the models produced which are based on data from past bridge projects. The main objective is to enable the user of the models to quickly and easily estimate the actual cost as well as the actual duration of a bridge using a model that will accept input values that will be known before the start of the project. A contractor that constructs bridges could use his past project data to adapt neural network models to his requirements.

The key factor for further development of the created models is the continuous enrichment of the database with new data. The larger the sample, the greater the reliability of the models. Also important for the development of models is the introduction of more independent variables in correlation analysis such as the conventional construction cost, the conventional construction duration, the weather conditions, the contractor's productivity, the contractor's equipment, the human resources, etc. The addition of new quantities and data to neural networks will result in the development of new models that could be compared to existing ones and important conclusions could be drawn.

The models for predicting actual cost and actual duration could be used by contractors, design firms, technical consultancies, local government agencies and technical tendering authorities in such a way as to be able to argue if a budget for the construction of a bridge and its construction schedule correspond to reality. In this way, they will be able to assess whether a discount given in a tender offer for the construction of a bridge can be achieved without affecting the quality of the project. Also based on a given construction budget, the contractor will be able to estimate the discount that could be given to the tender offer for the construction of a bridge. Besides, the construction schedule of a bridge could be predicted by the interested parties with the use of minimal data.

Regarding the actual cost of a bridge, the model that gave the lowest MSE is Model No 8 and the following apply to it:

- Quantitative + Qualitative + WEKA input variables (MSE = 3.84886e-05) includes the following independent variables: Budgeted cost, the quantity of deck concrete.

Regarding the actual duration of construction of a bridge, the model that gave the lowest MSE is Model No 33 and the following apply to it:

- Quantitative + Qualitative + WEKA input variables (MSE = 5.89463e-05) (the combination came from the WEKA application) includes the following independent variables: Budgeted cost, the quantity of deck concrete.

It is observed that both in the case where the dependent variable is the actual cost but also in the case where the dependent variable is the actual duration, the group of independent variables that gave the most efficient models with the least MSE are common and they are the budgeted cost and the quantity of deck concrete. It is recalled that in the case where the actual duration is the dependent variable, the combination of the resulting independent variables which gave the lowest MSE came from the subgroup identified by the WEKA application.

The smaller the number of independent variables that a model uses, the better it is for the user, because it makes the model simpler to use and also makes it easier to collect and record the required data.

The approach used to determine the most efficient variables and produce the models in this study could be adopted and applied to any other of construction project. The results from the models which were created seem to be very satisfactory and promising. The top model, for both dependent variable cases which were evaluated, contains the same two independent variables which are the budgeted cost and the quantity of deck concrete. The data of these variables can be easily determined in the early stages of a bridge technical study, even from the feasibility stage. Also, these variables contain low risk and are reliable. At the same time, the data of these variables are common to all bridge projects and this gives confidence that the proposed models could be successfully applied to other bridge projects. Of course, as the number of projects whose data are integrated into the neural network training database increases, so does the reliability of predictive models.

More research in this area of forecasting could be aimed at developing other methods and techniques and comparing their results with the results of the present study. Also, a very interesting comparison would be the one with relative results obtained from algorithmic models. On the other hand, the number of independent variables involved in the design of neural networks could increase and new optimal combinations of variables could emerge that lead to new neural network models. Along with the increase of variables, the number of projects included in the database could also increase. Finally, it is noted that in the present research all the effort for the production of neural network models was based on the ability of the FANN TOOL application to select and build the optimal neural network structure by the user's choice of the "Cascade" method. So, another suggestion for future work would be for users to experiment by designing their neural networks that might lead to better-performing models.

ACKNOWLEDGMENT

Parts of this work have been developed under a postgraduate thesis of the MSc in Construction Management, Hellenic Open University. Moreover, parts of this study have been associated with research conducted at the MSc in Project and Programme Management, University of Thessaly.

REFERENCES

[1] Al-Saadi, A.M., Zamiem, S.K. and Al-Jumaili, L.A. (2017) 'Estimating the Optimum Duration of Road Projects using Neural Network Model', *International Journal of Engineering and Technology (IJET)*, Vol. 9, No.5, pp.3458-3469

[2] Al-Zubaidi, E.D.A., Yas, A.H. and Abbas, H.F. (2019) 'Guess the time of implementation of residential construction projects using neural networks ANN', *Periodicals of Engineering and National Sciences*, Vol. 7, No. 3, pp.1218-1227.

[3] Antoine, A.L., Alleman, D. and Molenaar, K.R. (2018) 'Examination of project duration, project intensity, and timing of cost certainty in highway project delivery methods', *Journal of Management in Engineering*, Vol.35 No. 1, p. 04018049.

[4] Arafa, M. and Alqedra, M. (2011) 'Early Stage Cost Estimation of Buildings Construction Projects using Artificial Neural Networks', *Journal of Artificial Intelligence* 4 (1):63-75.

[5] Aretoulis, G.N. (2019) 'Neural network models for actual prediction in Greek public highway projects', *Int. J. Project Organisation and Management,* Vol. 11, No. 1, pp.41-64.

[6] Aretoulis, G.N., Angelides, D.C., Kalfakou, G.P., Fotiadis, G.S. and Anastasiadis, K.I. (2006) 'A Prototype System for the Prediction of Final Cost in Construction Projects', *Springer Operational Research, An International Journal*, Special Issue, 6(3), pp. 423-432.

[7] Aretoulis, G.N., Kalfakou, G.P. and Seridou, A.A. (2015) 'Project Managers' Profile Influence on Design and Implementation of Cost Monitoring and Control Systems for Construction Projects', *IGI-Global IJITPM*, 6(3), pp. 1-25

[8] Aziz, A.M.A. (2007). *Performance Analysis and Forecasting for WSDOT Highway Projects*, Washington State Transportation Commission, Agreement T2695 Task 92, Washington.

[9] Barros, L.B., Marcy, M. and Carvalho, M.T.M. (2018). Construction Cost Estimation of Brazilian Highways using Artificial Neural Networks. *International Journal of Structural and Civil Engineering Research*, vol 7, No 3, pp. 283-289.

[10] Chan D.W.M and Kumarswamy M.M. (1996). An evaluation of construction time performance in the building industry. *Elsevier Build. Environ*., 31(6), pp. 569-578.

[11] Chandanshive V.B. and Kambekar A.R. (2019) 'Estimation of Building Construction Cost Using Artificial Neural Networks', *Journal of Soft Computing in Civil Engineering*, 3-1, 91-107.

[12] Cheng, M.Y., Tsai, H.C. and Sudjono, E. (2010a) 'Conceptual cost estimates using evolutionary fuzzy hybrid neural network for projects in construction industry', *Expert Systems with Applications*, Vol. 37, No. 6, pp.4224–4231.

[13] Cheng, M.Y., Tsai, H.C. and Sudjono, E. (2010b) 'Evolutionary fuzzy hybrid neural network for project cash flow control', *Engineering Applications of Artificial Intelligence*, Vol. 23, No. 4, pp.604–613.

[14] ElSawy, I., Hosny, H. and Razek, M.A. (2011) *A Neural Network Model for Construction Projects Site Overhead Cost Estimating in Egypt*, αrXiv preprint arXiv: 1106.1570.

[15] FANN Tool Users Guide (online) https://www.slideshare.net/bluekid/fann-tool-usersguide

[16] Field, A. (2009) *Discovering Statistics Using SPSS*, 3rd ed., Sage

Publications, London.

[17] Gab-Allah, A.A., Ibrahim, A.H. and Hagras, O.A. (2015) ‘Predicting the construction duration of building projects using artificial neural networks’, *Int. J. Applied Management Science*, Vol.7, No.2.

[18] Glymis, E., Kanelakis, A., Aretoulis, G. and Mastoras, T. (2017) ‘Predicting highway projects’ actual duration using neural networks’, LC3 2017: Volume I – *Proceedings of the Joint Conference on Computing in Construction (JC3)*, Heraklion, Greece, pp.691–697 (online) https://doi.org/10.24928/JC3-2017/0260

[19] Guerrero, M.A., Villacampa, Y. and Montoyo, A. (2014) ‘Modeling construction time in Spanish building projects’, *International Journal of Project management*, Vol. 32, No.5, pp.861-873.

[20] Hodgson, D., Paton, S. and Cicmil, S. (2011) ‘Great expectations and hard times: the paradoxical experience of the engineer as project manager’, *International Journal of Project Management*, Vol. 29, No. 4, pp.374-382

[21] Hosseinian, S. and Reinschmidt, K.F. (2015) ‘Finding best model to forecast construction duration of road tunnels with new Austrian tunneling method using Bayesian inference: case study of Niayesh highway tunnel in Iran’, Transportation Research Record: *Journal of the Transportation Research Board*, Vol. 2522 No. 1, pp. 113-120.

[22] Irfan, M., Khurshid, B.M., Anastasopoulos, P., Labi, S. and Moavenzadeh, F. (2011) ‘Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type and contract type’, *International Journal of Project Management*, Vol. 29, No. 1, pp.78–92.

[23] Jiang, Y. and Wu, H. (2007) ‘A method for highway agency to estimate highway construction durations and set contract times’, *International Journal of Construction Education and Research*, Vol. 3, No. 3, pp.199–216.

[24] Juszczyk, M. (2017) ‘The challenges of nonparametric cost estimation of construction works with the use of artificial intelligence tools’, *Procedia Engineering,* Vol. 196, pp.415–422.

[25] Kang, H.W. and Kim, Y.S. (2018) ‘A model for risk cost and bidding price prediction based on risk information in plant construction projects’, *KSCE Journal of Civil Engineering*, pp.1–15, Article in press, DOI: 10.1007/s12205-018-0587-4.

[26] Koo, C., Hong, T. and Hyun, C. (2011) ‘The development of a construction cost prediction model with improved prediction capacity using the advanced CBR approach’, *Expert Systems with Applications*, Vol. 38, No. 7, pp.8597–8606.

[27] Liu, W. (2011), “Duration estimation method for highway construction work. Management and service science (MASS)”, *International Conference on, 2011*, IEEE, pp. 1-3.

[28] Marzoughi, F., Arthanari, T. and Askarany, D. (2018), “A decision support framework for estimating project duration under the impact of weather”, *Automation in Construction*, Vol. 87, pp. 287-296.

[29] Migliaccio, G.C. and Shrestha, P.P. (2009), “Analysis of design-build procurement activities durations for highway projects”, Construction Research Congress 2009: *Building a Sustainable Future*, pp. 229-238.

[30] Petruseva, S., Zujo, V. and Zileska-Pankovska, V. (2012) ‘Neural Network Prediction Model for Construction Project Duration’, *International Journal of Engineering Research and Technology*, Vol.1(02), ISSN.2778-0181.

[31] Roxas, C.L. and Ongpeng, J.M. (2014) ‘An artificial neural network approach to structural cost estimation of building projects in the Philippines’, *DLSU Research Congress*, March 6-8.

[32] Sonmez, R. (2011) ‘Range estimation of construction costs using neural networks with bootstrap prediction intervals’, *Expert Systems with Applications*, Vol. 38, No. 8, pp.9913–9917.

[33] Stoy, C., Pollalis, S. and Dursun, O. (2012) ‘A concept for developing construction element cost models for German residential building projects’, *International Journal of Project Organisation and Management*, Vol. 4, No. 1, pp.38–53.

[34] Titirla, M. and Aretoulis, G.N. (2019) ‘Neural network models for actual duration of Greek highway projects’, *Journal of Engineering Design and Technology*, Vol. 17, No. 6, pp. 1323-1339.

[35] Wang, Y.R., Yu, C.Y. and Chan, H.H. (2012b) ‘Predicting construction cost and schedule success using artificial neural networks ensemble and support vector machines classification models’, *International Journal of Project Management*, Vol. 30, No. 4, pp.470–478.

[36] WEKA (2018) (online) https://www.cs.waikato.ac.nz/ml/weka/

[37] Yadav, R., Vyas, M., Vyas, V., and Agrawal, S. (2016) ‘Cost Estimation Model for Residential Building using Artificial Neural Network’, *International Journal of Engineering Research and Technology*, Vol. 5, No. 2, pp. 312-314.