International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2015): 6.391

Late Payments for Contractors Working for Bahrain Government Building Construction Projects: Part II (Modelling Using Artificial Neural Networks and Regression)

Sahar K. A. Falamarzi[1], Saad M. A. Suliman[2]

[1, 2]Department of Mechanical Engineering, College of Engineering/ University of Bahrain, Bahrain

Abstract: ***The problem of late payment is considered one of the major issues in the construction industry, it is an important issue faced by many countries including Kingdom of Bahrain, and it has many consequences on building construction projects. The main objectives of this research is to create a prediction model to predict the payment delay in days for Interim payments (Model I) and variation payments (Model V) for Ministry of Works (MoW) building construction projects. The main factors of payment delay were identified in Part I of the research and used for the development of models using artificial neural network (ANN) and multi linear regression analysis (MLRA). Finally, the study compared the ANN approach to MLRA and concluded that the estimation accuracy of ANN approach is better than MLRA analysis for payment delay in (days) as it showed more promising results. The best ANN for Model I and Model V were found to be Model I-48 and Model V-60, respectively.***

Keywords: Interim payment delay, Variation payment delay, artificial neural network, multi linear regression analysis

1. Background

Late payments have a huge effect on many industries including construction industry. Late payments are also affecting the projects of the Construction Projects Directorate (CPD) in Ministry of Works (MoW) – Bahrain. This study aims to use the factors causing payment delay which were identified in Part I of this research to create prediction models to establish the interim payment delay in (days) and variation payment delay in (days) faced by the contractors working for MoW. This model helps in predicting the delay in payment and ensures smooth completion of work.

In the following sections the literature review will be presented first, where it includes related previous studies, followed by research methodology, and historical data collection and analysis for modelling. Then modelling with neural networks and regression. After that development of interim payment model (Model I) and variation payment model (Model V).Finally, conclusions will be presented and recommendations will be suggested in order to overcome the payment delay issue.

2. Literature Review

2.1 Introduction

Many previous studies discussed the Modeling of ANN and MLRN in Construction Project Management. Börner et al., (2012) stated that a proper description of a model is that it's a structural illustration of a certain object that has critical features to present the actual situation for a set of data in the visual, mathematical or as a computer simulation form. Waziri et al., (2014) stated that the most positive outcomes were found in the analysis done using regression and artificial intelligence for the building construction industry. These two methods showed high understanding of the elements affecting construction durations and their relationships.

2.2 Modeling with ANN and MLRN in Construction Project Management

In 2004 Bordat et al. (2004) analyzed and assessed the extent of cost overruns, time delays, and change orders problem associated with Indiana Department of Transportation construction projects. After identifying influential factors, regression models were developed. The models were created using 2000 randomly selected projects to estimate values of cost overruns, time delay and change orders for future projects. The model provided information about the significant factors affecting cost overruns, time delays, and change orders, which are: bid amount, project type, location by district, weather, and bid comparison variables.

Ayman et al. (2008) study was about risk prediction, which was inducted in Jordan by proposing a probabilistic model to predict the risk effects on time and cost of construction projects. Statistical regression model was developed using real data of 140 projects, it estimated project cost and duration based on historical data. Customized multiple regression models were developed for each project type to obtain statistically reliable results. In conclusion, the proposed model predicted the project cost and duration with a precision of ±0.035%.Odabaşi (2009) conducted a study on estimating construction duration based on many factors. Simple Linear Regression (SLR) and Multiple Linear Regression (MLR) analyses were conducted by investigating the influencing factors, using previous case study buildings, to predict the construction duration of a project. The confidence in estimation of the regression analyses was investigated, and finally an MLR model was obtained which was based on two parameters: the area of the building, and the area of its front. As a conclusion and contrasting to previous studies, no significant effect was noticed for the cost on project duration.

Yahia et al. (2011) developed an ANN model to predict time contingency in construction projects by identifying the important factors based on a comprehensive survey filled by the Egyptian construction experts. The model was developed in order to have a more reliable prediction for the amount of time contingency that should be added to the scheduled completion time by project planner. **Also** in 2011 an ANN model was developed by Elsawy et al. (2011) for estimating site overhead cost using 52 actual real-life cases of building projects, constructed in Egypt, as training materials. The model presented the site overhead costs as a percentage from the total project cost.

In 2015 ANN model was developed by Naik and Radhika (2015) for the estimation of cost and duration for highway road construction projects. The database was collected from previous projects, normalized and then used as inputs and targets for the ANN models development. The models are trained, tested and validated using MATLAB R2013a Software. The model was trained with feed forward back propagation learning algorithm. The performance evaluation of the ANN was done using MAPE (Mean Absolute Percent Error) by comparing the output values from the ANN with the actual values. The best results were given by the ANN with training function trainlm which is a function that uses Levenberg-Marquardt algorithm and (Nftool approach) with 2 layers and a hidden layer of 10 neurons.

Parminder (2016) developed an ANN model for predicting the time duration of construction projects in India, by first selecting the important factors through extensive literature study. A multi-layers model with back propagation learning algorithm was developed with several cases. The best model obtained was with one hidden layer containing 20 neurons, and with minimum root mean square error of 0.9845. The proposed ANN model was compared with other project planning techniques where it has found to be the most accurate and reliable tool for project duration prediction.In 2016 an ANN was used by Mensah et al. (2016) to predict the duration of prefabricated steel bridge projects in Ghana. Data for 18 completed bridge construction projects were collected to get the independent items to be used as an as input variables and the actual durations as output variables. The model was developed with a feed forward back propagation algorithm and the number of neurons in the hidden layer was obtained by trial and error. The Accurate results of the model was obtained with a coefficient of determination (R2 =0.998) and MAPE of 4.05%. The study has shown that the developed model is suitable for estimating the duration of a bridge project.

In 2017 a model was developed by Renuka et al. (2017) to estimate the expected percentage of the time overrun for a particular construction project during the planning stage. Six major groups causing projects time delay were identified. A questionnaire was answered by project managers by asking them to fill the percentage of delay for each group related issues and the overall delay duration of the construction projects. Thereafter, a regression model was used to analyze the collected data, and the results of the model were used to study the relationship between the percentage of delay and the major group related issues that causes delay.

It was found that most of the previous researches did not include modeling of payment delay and mainly discussed and identified the causes and effect of this issue. Moreover, no research was found to study the payment delay in Kingdom of Bahrain and precisely in governmental building construction projects. Therefore, this study is aimed to identify the causes and effect of payment delay in governmental construction projects in Bahrain and to model the delay in payment based on real case studies in this field.

3. Research Methodology

3.1 Introduction

This section comprises the techniques undertaken to develope a predictive payment model to predict the delay in two type of payments practiced in MoW projects. These payments are: interim payments and variation payments.

3.2 Development of the Model

Development of the model was achieved through the following steps:

1-Once the significant factors are identified from the distributed Questionnaire as explained in Part I of this research, the input factors for the models are identified according to the availability of data and compatibility with construction project directorate (CPD) cases.
2-The input data for the significant factors are collected from CPD real case studies and is used for the model's development as independent variables (Input) to predict the dependent variable (Output), which is the payment delay (days).
3-Model verification is carried out using some of the collected data.

These models support the study by predicting the payment delay for governmental building construction projects and giving a better understanding for the actual duration needed for a payment to be approved and actually received from the employer. This increases the attention of this issue to overcome the payment delay when encountered without affecting the contractors, consultant or the progress of the project.

4. Historical Data Collection and Analysis for Modeling

4.1 Introduction

This section presents the data collection and analysis to predict the payment delay in days for interim payment delay (Model I) and variation payment delay (Model V) using historical payment cases for building construction projects accomplished in Bahrain during years (2014-2016). The independent and dependent variables used for the design of the models are presented in this chapter along withtheir relevant data, which are being collected from real case projects.

4.2 Independent and Dependent Variables

The independent variables (input) used in the models design for payment delay prediction are expressed in terms of the payment delay factors collected from the questionnaire results. The payment delay factors considered for Model I and Model V are chosen based on the top twenty factors ranked in Part I of this research by the consultants and contractors as shown in Table (4-1). Eleven of the factors are common in the two rankings of the consultants and contractors, thus, a total of 29 factors are analyzed for inclusion in the models to be developed.

1- "Changes in scope of work": This is considered to be the main reason for variation orders and extension of time in governmental projects. It is ranked in first place as per the contractors and consultants perspectives. This factor is affecting and causing many other factors where "change in cope"results in the occurrence of "Incomplete documents for variation claims by the contactor", "Delay in submitting valuation/claim by the contractor", "Delay in extension of time approval by the consultant", "Submission of claims with calculation mistakes by the contractor"and "Slow processing of variations orders by the consultant". This factor will be expressed indirectly in Model I by the factors: "Occurrence of variations orders"and "Occurrence of extension of time". While will be expressed in Model V by the factor: "Processing of variation orders".
2- "Incomplete documents for variation claims by the contactor": Variation orders with incomplete documents are rejected by the consultants and the contractor is requested to resubmit it along with complete attachments. The number of the occurrences of this factor is not available for the used historical cases, thus, it will not be considered for modeling.
3- "Delay in submitting valuation/claim by the contractor": This factor will be used in Model I to represent the number of days delayed by the contractor to raise his valuation (request for interim payment). While for Model V no delay occurred in the used historical cases by the contractor in submitting the claim (request for variation order). This is justified by the additional amounts included in each of the variations orders, which can provide an extra profit for the contractor.
4- "Delay in extension of time approval by the consultant": The number of delayed days for extension of time approval is not available for the used historical cases; therefore this factor is expressed by the "occurrence of extension of time "in each of the cases by the binary expression.
5- "Submission of claims with calculation mistakes by the contractor": Claims with calculation mistakes are rejected by the consultant and sent back to the contractor for rectifying and resubmission; No data are found for this factor, thus, it will not be considered for modelling.
6- "Slow processing of variations orders by the consultant": This factor will be used in the development of Model V and is expressed by the number of days required by the consultant to evaluate and approve the variation order submitted by the contractor.
7- "Delay in the progress of works and activities"and "In-appropriate implementation of projects program"are both expressed by the "Delay in progress of work "in days. This factor will be used in the development of Model I and Model V.
8- "Ministry of finance process"is noticed from the historical case studies to have a major effect on payment delay. Moreover, it is ranked in the 12^{th} position as per the consultants perspective; therefore, this factor will be used in Model I and Model V, and is expressed by the number of days needed for the owner to pay the amounts due to the contractor.
9- "Delay in valuation review and evaluation process"and "Delay in issuance of payment certificate process" are ranked in the 4^{th} and 8^{th} position, respectively, from contractors perspective. These factors will be used in Model I and Model V, and are both expressed by the total number of days required for the valuation to be reviewed, approved and issued in the form of payment certificate.
10- "Lack of periodical meetings to address payment problem":This factor occurs in all the cases, and it is evidence of occurrence is availability of minutes of meetings of each periodical meeting. It is not included in the models since it is positively occurring in all used historical cases.
11- The number of years of experience for each of the contractor and consultant are ranked in 8^{th} and 18^{th} position, respectively, as per the consultants and contractors perspective, respectively. These factors will be used in both Model I and Model V.
12- "Long and bureaucratic process in governmental departments", "Inability to follow certain procedures of MoW", "Failure to understand the contract agreement", "Lack of decision making during construction", "Inaccurate bill of quantity"and

"Accuracy of payment scheduling program"are intangible factors, therefore they will not be included in the models. Although these factors have huge impact on the payment process since they can extend the period of the payment process and makes it longer, more complicated, and less flexible. On the other hand, the positive occurrence of these six factors can hugely improve the payment process and decrease the amount of delay.

13- Information on "Delays of documentation required for fulfilling payments", "Inadequate financial resource" and "Slow processing of final accounts" are not available in for the used historical cases, therefore, these factors will not be included in the models.

14- "Changes in rules and regulations", "Economic changes "and "Unavailability of funds "do not occur during the used historical cases. All the used historical cases were in years 2014-2016, where no major changes existed with regard to these factor to affect payment processes.

15- "Refusal to pay interest on late payment": this factor is not applicable in MoW projects because it is against civil law as discussed earlier in the questionnaire results and literature.

16- "Duplication of work":The frequency of work duplication is not available in the historical cases, therefore this factor is eliminated.

Table 4-1: Factors causing payment delay

Rank	Consultant perspective		Contractor perspective	
	Factors causing payment delay	Data Remarks	Factors causing payment delay	Data Remarks
1	Changes in scope	available	Changes in scope	available
2	Incomplete documents for variation claims.	unavailable	Long and bureaucratic process	Intangible factor
3	Long and bureaucratic process	Intangible factor	Delay in Extension of time approval	unavailable
4	Delay in the progress of works.	available	Delay in valuation review and evaluation.	available
5	Delay in submitting the payment valuation/Claim by the contractor.	available	Slow processing of variations orders.	available
6	Delay in Extension of time approval	unavailable	Slow processing of final accounts	unavailable
7	Submission of claims with calculation mistakes.	unavailable	Refusal to pay interest on late payment	inapplicable
8	Contractor's experience in governmental projects.	available	Delay in issuance of payment certificate Process.	available
9	Inappropriate implementation of projects program.	available	Lack of decision making during construction	Intangible factor
10	Delays of documentation required to fulfill payments.	unavailable	Inaccurate bill of quantities	Intangible factor
11	Inability to follow MoW procedures.	Intangible factor	Inappropriate implementation of projects program.	available
12	Ministry of Finance processes	available	Changes in rules and regulations	Didn't occur
13	Failure to understand the contract agreement	Intangible factor	Economic changes	Didn't occur
14	Accuracy of estimation	Intangible factor	Unavailability of funds	Didn't occur
15	Slow processing of final accounts	unavailable	Incomplete documents for variation claims.	unavailable
16	Slow processing of variations orders.	available	Delays of documentation required to fulfill payments.	unavailable
17	Lack of decision making during construction	Intangible factor	Lack of periodical meetings to address payment problems	Exists in all cases
18	Accuracy of payment scheduling program	Intangible factor	Consultant's experience	available
19	Inadequate Financial resource	unavailable	Duplication of work	unavailable
20	Lack of periodical meetings to address payment problems.	Exists in all cases	Submission of claims with calculation mistakes.	unavailable

The 11 factors shown in Table (4-2) and the 10 factors shown in Table (4-3) are chosen as the input for Model I and Model V, respectively. The rest of the top twenty factors of Table (4-1) are not included in the modelling for reasons like: unavailability of data or non-occurrence of the factor in the real historical cases used to build the model as shown in the remarks column in Table (4-1). Moreover, additional factors are added to the models as per the experts' recommendations, which include the following:

a) "Amount paid in Bahraini Dinar": This factor is used in Model I and Model V for each interim payment certificate or variation payment.

b) "Total available contingency in the contract": This factor reflects the total amount of contingency available at the beginning of each project, and it is usually 10% of the project budget. This factor is used for Model V development because it is concerned with variation orders.

c) "Balance of the contingency after each variation payments": This factor is used in Model V, it is shown in the historical data that with the decrease of the contingency amount, approval process for variation orders becomes longer.
d) "Type of funding resource": There are two types of funding governmental building projects: Local and external, with different payment approval process for each. This factor is used in Model I only because all the historical cases used for Model V development are externally funded.

Table 4-2: Interim payment factors used for model I

	Factors	Value
1	Delay in the submission of payment evaluation by the contractor	Days
2	Valuation review	Days
3	Payment processed by owner	Days
4	Variation orders occurrence	Binary expression
5	Extension of time occurrence	Binary expression
6	Payment amount	Bahraini Dinar.
7	Contractor experience	Years
8	Consultant experience	Years
9	Externally funded projects	Binary expression
10	Locally funded projects by government of Bahrain.	Binary expression
11	Work progress	Days

Table 4-3: Variation order payment factors used for Model V

	Factors	Value
1	Variation orders evaluation and approval	Days
2	Claim review	Days
3	Payment processed by employer	Days
4	Allocated contingency in the contract	Bahraini Dinar
5	Balance of contingency	Bahraini Dinar
6	Extension of time occurrence	Binary expression
7	Payment amount	Bahraini Dinar.
8	Contractor experience	Years
9	Consultant experience	Years
10	Work progress	Days

4.2 Data Collection and Types

The payment delay models are designed using historical data for actual governmental projects. One hundred and fourteen interim payment cases for building projects are used for the development of Model I, while 91 variation payment cases for building projects are used for Model V development. The chosen cases are all for school projects constructed in the years between 2014-2016. The values of the models' factors are extracted from the cases and organized in an Excel sheet for the design purpose. The values of the payment delay factors which are functioning as the input of the models design are defined in two ways. First, as a numerical input, and secondly as binary input (0, 1), where (1) expresses the positive reaction as the availability of the factor in each case and (0) expresses the negative reaction like unavailability of the factor in each of the payment cases. After defining the input factors, the historical data needed for the development of Model I and Model V are expressed as follows:

4.2.1 Factors used in interim payment delay model (Model I)

For interim payment delay model development, eleven factors are chosen and expressed as shown in Table (4-2). They are illustrated below:

1. Delay in submitting valuation by the contractor: This factor is expressed by the number of days the submission of payment notice is delayed by the contractor at the end of each month as per payment schedule.
2. Claim review by the consultant: This factor is expressed by the total number of days of evaluation and issuance of payment certificate by summing the following: -
3. Total number of days from receiving the claim by the project manager until it is forwarded to Cost Engineering Department (CED).
4. The total number of days spent by CED for the evaluation and issuance of payment certificate and forwarding it to Finance Resources Directorate (FRD).
5. Total number of days for FRD to check and forward payment certificate toMinistry of Finance (MoF) in case of governmental funded projects and to the funding entity in case of externally funded projects.
6. Payment processing by owner: This factor is expressed by the number of days spent by the owner to make full payment to the contractor.
7. Variations orders occurrence: This factor is expressed by binary expression: 1 in case of the existence of variation within each payment case, and by 0 if there is not.

8. Extension of time occurrence: This factor is expressed by binary expression: 1 in case the payment is affected by extension of time based on expert's opinion and by 0 in case it is not.
9. Payment amount: It is expressed by the due amount of payment certificate in Bahraini Dinar.
10. Contractor experience: This factor is expressed by the total number of years of experience of the contractor's engineer who process the claim.
11. Consultant experience: This factor is expressed by the total number of years of experience of the consultant's (MoW) engineer who process the claim.
12. Externally funded projects: This factor is expressed by binary expression: 1 in case it is an externally funded project, and by 0 in case it is not.
13. Locally funded projects by government of Bahrain: This factor is expressed by binary expression: 1 in case it is locally funded project, and by 0 in case it is not.
14. Work Progress: It is the total number of time delay of the project at the time of each payment.

The actual payment delay in days (dependent variable) for Model I is counted by investigating the following factors:

1. The delay in submitting payment valuation by the contractor: This factor is calculated by subtracting the actual date of submitting the payment valuation by the contractor from the planned date of submitting the payment valuation by the contractor.
2. Number of day for payment claim to be reviewed and approved by consultant: This factor is calculated by subtracting the date of claim final approval by MoW from the date of claim receival from the contractor.
3. Number of days for payment to be paid by owner: This factor is calculated by subtracting the date of payment deposit in contractor bank account from the date of payment final approval is received from MoW.

4.2.2 Factors used in variation payment delay model (Model V)

The variation payments process in MoW as explained in Part I of this research is a long process, and different than the interim payment process especially for externally funded projects, where it involves three stages: prior approval of variation from funding entity, formal approval of variation from funding entity, and finally certifying the payment. While for locally funded projects it includes only two stages: variation order approval, and payment certification stage. Table (4-3) shows the factors used in model V and their expressions as illustrated below:

1. Variation evaluation and approval: This factor is expressed as follows:
   a) For locally funded projects, is given by the total number of days for variation orders approval.
   b) For external funded projects, is given by the total number of days for prior and formal approval of variation.
2. Claim review: as defined in Subsection 4.2.1.
3. Payment process by owner: as defined in Subsection 4.2.1.
4. Allocated contingency in the contract: It reflects the total amount of contingency (Bahraini Dinar) allocated for the project.
5. Balance of contingency: This factor is expressed by the amount of contingency (BD) balance in the contract at the time of the variation approval. The collected data has shown the effect of this factor on the approval decision of both the consultant and the client for each variation case, where less number of variations are approved with the decrease of contingency balance during construction, where only the most important variation items are approved.
6. Extension of time occurrence: as defined in Subsection 4.2.1.
7. Payment Amount: It is expressed by the amount of variation claimed in Bahraini Dinar.
8. Contractor Experience: as defined in Subsection 4.2.1.
9. Consultant Experience: as defined in Subsection 4.2.1.
10. Work Progress: as defined in Subsection 4.2.1.

The actual payment delay in days (dependent variable) for Model V is calculated by investigating the following factors:

1. Number of days for Variation order evaluation and approval: This factor is calculated by subtracting the date of variation order final approval by consultant from the date of submission of variation order request by the contractor.
2. Number of day for payment claim to be review and approved by consultant: as defined in Subsection 4.2.1.
3. Number of days for variation payment to be paid by owner: as defined in Subsection 4.2.1.

5. Modelling with Neural Networks and Regression

5.1 Introduction

In this study two types of models are applied for the prediction of interim and variation payment delay which are: artificial neural network model and regression model. Each of the applied models has different software tools, architecture and features. These differences are explained further in order to get an assumption of the most suitable model for the prediction of interim and variation payments delay. Below is an explanation of the modeling techniques used in this study:

5.2 Modeling Techniques

The first modeling techniques used is the artificial neural network with multiple layers: are input layer, one or more hidden layers, and one output layer (Goyal and Goyal, 2012). Three learning algorithms are used in this study: Levenberg-Marquardt algorithm, Bayesian regularization algorithm and Scaled Conjugate gradient.The Forecasting of the output is performed through 3 stages which are training, validation and testing. To perform the training; the learning step starts by discovering linear relationships between the inputs and the output data by assigning weight values to the links between the neurons. The second step is feed forward of the network, where the neurons are added to the hidden layer and the input values in the first layer are multiplied by the weights and passed to the second (hidden) layer to create one iteration (Al-Sobiei et al., 2005). At each iteration, the weights are adjustedto minimize the calculated error measure between the output produced and the targeted output. The error minimization process is repeated until an acceptable result is reached (Iranmanesh and Zarezadeh, 2008).In case the predicted results do not meet the desired output and this is measured by the validation stage where the performance (error) is calculated, then the back propagation of error signals takes place in which error is propagated back to all the elements in the prior layer and finally updating of the weights and biases based on the error signal (Amita et.al, 2015). The testing stage is finally performed by measuring the performance of the model using an independent set of data that is not included in the training and validation stage.

The neural models in this study are developed using Neural Networks Toolboxby MATLAB. In MATLAB there are several neural networks tools that help the user to perform any kind of neural networks smoothly and easily such as using the Neural Network Fitting Tool (nftool) and (nntool), as described in "Graphical User Interface "(Demuth et al., 2009).

The second modeling technique used is MRA. Microsoft Excel (Data analysis toolpak) version 2016 is used to develop regression models. Data analysis tool by Excel is used to develop different engineering analyses. The Data Analysis ToolPak is an Excel add-in tool. It contains more widespread functions, including some useful inferential statistical tests for example regression that can provide single and multiple linear regression (Rose et al., 2014).

The models that are used in this study are presented in Figure (5-1) and are tested for both Model I and Model V. Both MATLAB programme and Microsoft Excel are used for the development of the neural network and regression model, respectively.

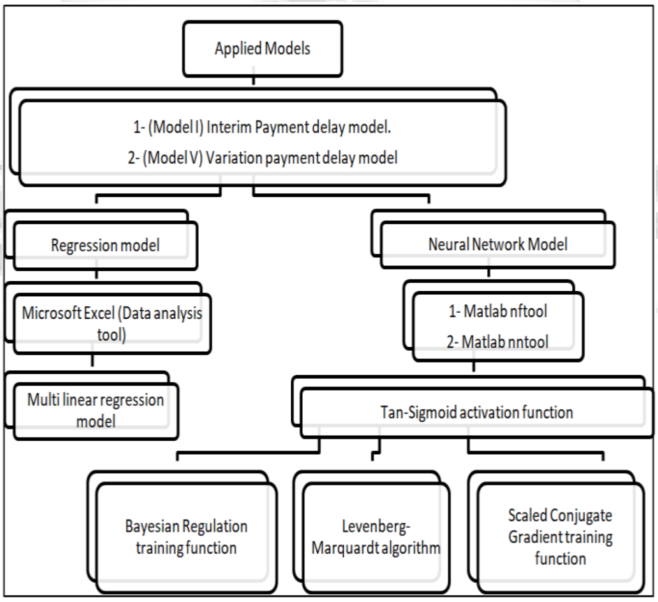

Figure 5-1: Approaches adopted for models development

5.3 Design of Models

Two models are developed using ANN Toolbox by MATLAB, while the preparation of the input data and statistical computations had been performed by applying Microsoft Excel. The first model is build based on historical data for 114 interim payments data points taken from real governmental building construction projects, while the second model is build based on historical data for 91 variation orders payments taken from other governmental building construction projects. The development of the models includes the steps shown in Figure (5-2) starting from identifying the input factors and collecting historical data as defined previously in Section (4.2), model development, calculation of the performance and finally comparision of different resullts.

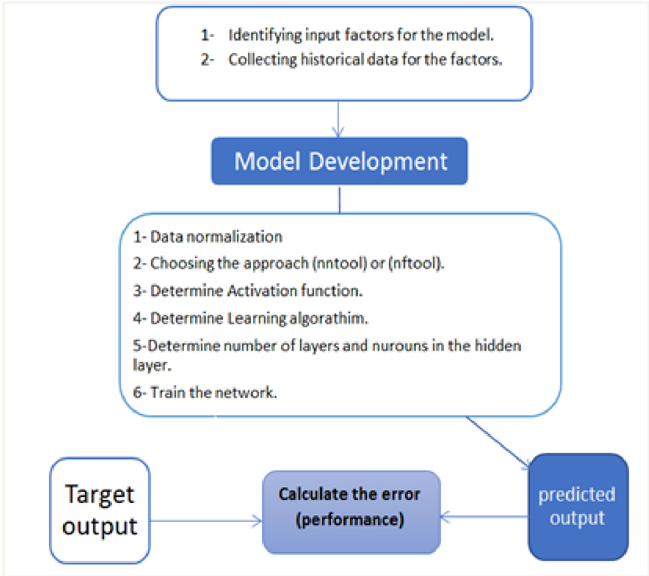

Figure 5-2: ANN Models Development

The Payment delay models are developed using the 114 actual interim payment cases for Model I and 91 actual variation payment cases for Model V. The data used in building each of the two models are divided into 3 subsets, (70%) used for training the network and (15%) for validation, and (15%) for testing. The program accepts up to two layers including 1 hidden layer and 1 output layer which are approved to be enough for most engineering problems. Multi-layers ANN models are developed and compared with each other using 2 approaches: 1) nftool (Neural Network Fitting Tool); and 2) nntool (in Network/Data Manager window) from MATLAB software. Hit and trial method is used to test different combinations of training algorithms and hidden neurons to train the models. Three chosen algorithms are used: Bayesian regularization, Levenberg Marquardt algorithm and Scaled conjugate algorithm. These algorithms are all tested along with a tan-sigmoid activation functionto work best with the normalized factors of a range [-1, 1]. In this study the frequently used "backpropagation" network was implemented.

The significant factors used in the development of the regression models for interim payment delay (Model I) and variation payment delay model (Model V) are the same as the factors used in the development of the neural model as defined in Section (4.2).These factors are working as the independent variables for the multiple regression models created by using Microsoft Excel (Data analysis tool).

The performance of the models was recognized through comparing their predicted output over a test sample. The test sample consists of dataset of 18 historical interim payment cases for Model I and 14 historical variation payments cases for Model V. The evaluation approaches that have been used to test the models are as follows: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Coefficient of Determination (R squared values).

6. Development of Interim Payment Delay Model (Model I)

6.1 Introduction

A total of thirty-three ANN models are trained, validated and tested for model I with different number of hidden neurons and training algorithms using MATLAB nntool. The one hundred and fourteen cases are divided into three sets: the training set consisting of 79 cases; the validation set consisting of 17 cases; and the testing set consisting of 18 cases. The below subsections show the performance of the test samples for multilayer (2 layers) feed-forward network with tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer for predicting interim payment delay. The training algorithms used are Levenberg-Marquardt algorithm, Bayesian Regulation and Scaled Conjugate Gradient. The performance of the training and validation samples for the thirty-three trial models are not involved in the evaluation of models to eliminate the risk of misevaluation, because during the training and learning some models may accomplish good results during the training and validation data set, but very poor on new data which is the test set. Therefore, the evaluation of the models' performance is done by comparing their predicted output over an independent test sample as mentioned in Section 5.3.

6.2 Developing Model I Using MATLAB nntool

A total of thirty-three ANN models are trained, validated and tested for model I with different number of hidden neurons and training algorithms using MATLAB nntool. The performance of the test samples for multilayer (2 layers) feed-forward network with tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer for predicting interim payment delay are compared as shown in Tables (6-1), (6-2) and (6-3). The training algorithms used are Levenberg-Marquardt algorithm, Bayesian Regulation and Scaled Conjugate Gradient. By comparing the computed results for each of the generated models using MATLAB (nntool), Model I-12 with 1 neuron in hidden layer based on Bayesian Regulation algorithm and tan-sigmoid function has been chosen. It has the lowest error values for Interim payment prediction with RMSE of 11.370, MAE of 8.947and MAPE value of 23.266% as shown in Table (6-4). The best performance for the training sample of this model is at 29 epochs with MSE of 0.003, and R value of 0.999. The validation sample performance is not available because the Matlab Programme sets the validation parameter for Bayesian Regulation as zero.

Table 6-1: Performance of test sample Model I using (nntool and Levenberg-Marquardt)

Target Output	Model no										
	I-1	I-2	I-3	I-4	I-5	I-6	I-7	I-8	I-9	I-10	I-11
24	93.926	16.347	18.564	47.799	84.902	9.000	9.587	23.127	9.290	18.129	9.030
64	108.677	70.601	22.939	90.319	23.740	9.000	47.781	61.102	46.180	81.232	10.813
115	110.088	89.622	21.427	101.665	81.242	9.000	110.794	92.930	111.864	67.441	12.330
26	83.865	18.498	11.685	22.032	56.276	9.000	9.787	16.891	9.234	16.818	9.083
9	73.282	15.401	9.351	15.665	71.975	9.000	9.427	15.090	9.040	17.718	9.021
62	99.628	30.891	15.924	27.368	59.572	9.000	19.359	39.711	72.656	28.263	9.298
23	83.656	14.148	9.662	22.776	20.156	9.000	9.843	13.769	9.657	17.860	9.171
30	92.402	17.910	11.388	46.329	42.942	9.002	12.262	17.058	10.964	19.864	9.162
21	66.304	12.149	9.095	15.370	13.162	9.000	9.304	11.876	9.143	14.054	9.100
67	111.414	85.159	38.883	95.597	70.507	115.000	41.300	85.708	114.942	22.385	10.651
19	81.262	16.529	9.486	32.908	15.742	42.112	9.037	11.526	9.469	16.128	9.072
60	100.861	37.379	21.576	81.598	46.593	9.000	26.135	26.630	22.316	37.781	9.846
19	85.120	16.330	11.005	23.567	21.826	9.000	10.057	18.099	15.728	40.426	9.185
90	111.703	99.130	73.742	79.411	114.824	9.000	114.522	97.386	86.420	93.317	11.746
27	82.551	16.895	9.884	37.069	12.130	102.353	9.081	15.074	10.148	11.927	9.100
40	104.370	34.955	13.028	28.716	46.556	9.011	12.229	29.401	34.874	61.210	9.277
25	85.087	16.934	16.915	75.928	25.854	9.703	9.063	28.015	9.511	33.411	11.083
30	91.991	16.283	15.627	41.114	13.617	9.044	9.016	25.425	9.300	31.308	9.503
hidden neurons	1	5	10	20	30	40	50	60	70	80	100
RMSE	218.027	13.810	31.130	20.4822	26.828	45.812	20.506	13.557	18.902	20.792	41.708
MAE	51.389	11.468	22.862	16.30862	18.928	36.039	17.906	10.699	14.863	15.832	31.921
MAPE	205.458	31.410	50.304	48.51195	81.979	82.667	49.334	29.252	42.434	41.700	66.498

Table 6-2: Performance of test sample Model I using (nntool andBayesian Algorithm)

Target Output	Model no										
	I-12	I-13	I-14	I-15	I-16	I-17	I-18	I-19	I-20	I-21	I-22
24	25.050	17.477	17.475	22.959	21.383	20.648	22.289	16.727	22.282	23.640	22.437
64	76.232	71.928	72.772	85.897	79.351	107.301	74.319	93.546	81.607	74.358	78.327
115	89.302	102.356	100.520	104.348	99.893	98.212	99.983	102.968	100.353	98.495	100.758
26	19.825	15.884	14.868	17.636	18.606	13.133	20.411	14.338	18.297	19.976	19.830
9	15.448	12.559	12.067	11.584	14.447	10.147	14.753	11.071	14.115	15.385	15.146
62	42.727	50.226	35.396	52.817	39.717	58.123	52.465	47.741	38.248	40.325	42.057
23	19.643	17.461	15.024	15.086	19.186	17.315	19.443	14.671	16.765	19.287	20.024
30	24.349	27.808	18.937	20.667	25.132	19.253	24.373	19.959	20.138	23.407	27.859
21	14.585	12.032	11.726	10.926	14.460	12.777	12.906	10.843	12.419	14.435	14.773
67	89.108	110.309	103.360	113.464	99.467	111.220	111.275	109.050	97.999	97.019	100.660
19	16.325	12.340	13.649	12.529	18.061	11.350	13.976	11.310	13.569	16.808	18.573
60	45.082	49.173	41.040	58.139	45.657	55.474	57.677	52.892	42.640	42.018	48.129
19	20.027	15.499	15.101	15.363	20.781	16.564	18.023	14.409	16.863	20.172	21.849
90	93.144	110.842	110.099	111.836	106.787	113.813	113.157	111.676	109.472	104.770	108.350
27	18.219	16.023	15.805	15.196	19.733	15.354	17.236	13.349	15.533	19.069	20.350
40	50.062	52.785	35.539	59.895	42.697	65.444	62.444	45.880	41.659	43.907	44.993
25	19.217	14.468	17.203	17.662	24.776	19.428	14.807	14.547	17.832	20.655	26.631
30	23.748	15.021	18.524	19.716	26.168	14.291	18.746	16.065	21.967	24.160	26.703
hidden neurons	1	5	10	20	30	40	50	60	70	80	100
RMSE	11.370	14.410	14.737	15.626	12.367	18.609	14.828	15.953	13.566	12.035	12.106
MAE	8.947	11.314	12.138	11.702	9.098	13.722	10.812	12.911	11.052	9.241	8.748
MAPE	23.266	30.338	31.552	30.222	21.641	33.271	28.205	34.039	28.431	23.076	20.971

Paper ID: ART2018737 DOI: 10.21275/ART2018737

Table 6-3: Performance of test sample Model I using (nntool and Scaled Conjugate Gradient Algorithm)

Target Output	Model no										
	I-23	I-24	I-25	I-26	I-27	I-28	I-29	I-30	I-31	I-32	I-33
24	24.468	38.676	72.700	13.371	9.000	13.261	19.457	16.430	13.271	15.700	17.107
64	24.468	67.849	79.639	65.637	9.000	54.326	20.435	25.524	73.443	9.008	77.222
115	24.441	29.830	73.966	45.037	9.000	95.218	59.256	85.607	82.595	9.073	91.380
26	24.468	27.826	63.047	13.780	9.000	12.923	15.316	19.672	12.213	13.580	18.825
9	24.468	44.792	43.650	12.521	9.000	12.344	14.499	16.513	10.895	12.654	12.634
62	24.464	22.396	65.081	11.535	9.000	42.805	19.218	15.063	54.661	9.756	23.561
23	24.467	22.931	38.960	14.451	9.000	13.030	20.867	36.970	15.396	9.007	24.668
30	24.409	26.903	52.285	18.403	9.000	12.115	15.030	42.269	28.508	9.011	61.278
21	24.468	21.030	24.811	12.171	9.000	11.640	17.373	29.075	12.404	9.005	12.622
67	13.515	81.315	101.656	10.623	9.000	99.980	18.569	34.116	91.751	15.688	92.479
19	24.169	20.615	20.484	10.897	9.000	14.642	12.318	13.559	10.214	9.002	14.976
60	24.465	20.554	69.885	27.818	9.000	23.020	21.688	41.135	27.797	9.007	73.410
19	24.467	22.269	48.151	12.510	9.000	14.448	13.371	11.348	24.344	9.043	23.409
90	24.468	27.047	109.669	93.571	9.000	94.829	33.885	42.551	103.444	25.854	101.353
27	24.009	19.234	45.916	11.737	9.000	28.766	16.715	13.587	10.531	9.002	17.842
40	24.435	58.476	76.284	11.640	9.000	24.962	15.611	18.608	67.143	9.954	53.418
25	14.574	12.936	20.646	39.263	9.000	9.737	20.137	33.524	9.922	9.008	21.831
30	18.609	29.059	34.882	18.270	9.000	10.194	17.247	39.251	10.130	9.008	12.698
hidden neurons	1	5	10	20	30	40	50	60	70	80	100
RMSE	33.685	30.397	25.716	27.606	42.865	16.717	29.029	23.196	17.000	40.456	16.568
MAE	22.288	19.164	21.194	19.653	32.722	13.811	21.723	18.633	14.243	30.886	13.113
MAPE	45.545	51.649	77.948	46.400	68.027	37.630	44.857	45.060	38.214	65.931	33.694

Table 6-4: Selected Models for Model I Using Matlab (nntool)

Model No	MAE	RMSE	MAPE (%)	Training Algorithm
Model I-8	13.557	10.699	29.252	Levenberg
Model I-12	11.370	8.947	23.266	Bayesian
Model I-33	16.567	13.113	33.694	Scaled

6.2 Developing Model I Using MATLAB nftool

A total of thirty-six ANN models are trained validated and tested for model I using MATLAB nftool. The performance of the test samples for multilayer (2 layers) feed-forward network with tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer for predicting interim payment delay is compared as shown in Table (6-5), (6-6), and (6-7). The training algorithms used are Levenberg-Marquardt algorithm, Bayesian Regulation and Scaled Conjugate Gradient. By comparing the computed results for each of the selected models using MATLAB (nftool), Model I-48 with 10 neurons in hidden layer based onBayesian regulation algorithm and tan-sigmoid function has been chosen as it has the lowest error values for Interim payment prediction with RMSE of 1.520, MAE of 0.969 and MAPE value of 3.767% as shown in Table (6-8). The best performance for the training sample of this model is at 120 epochs with MSE of 1.457, and R value of 0.999 which is reflecting a good fit of results and high predictive power of the network.

Table 6-5: Performance of test sample Model I using (nftool and Levenberg-Marquardt)

Target Output	Model no										
	I-34	I-35	I-36	I-37	I-38	I-39	I-40	I-41	I-42	I-43	I-44
24	22.018	23.048	22.650	24.206	24.001	30.422	18.560	29.923	24.000	20.883	25.316
64	60.013	62.369	60.456	64.012	61.526	62.663	64.055	64.109	42.801	62.648	59.724
115	111.773	113.039	112.733	115.049	115.001	118.119	125.750	115.176	80.798	81.705	106.993
26	26.713	27.564	27.781	26.205	26.001	30.942	25.994	26.138	30.271	9.556	17.512
9	11.526	12.253	11.827	9.257	9.001	14.780	9.019	9.122	9.000	5.697	10.802
62	57.748	58.260	56.397	62.055	62.001	71.629	62.020	82.456	62.000	54.458	60.729
23	28.245	27.862	30.198	28.467	23.979	32.895	32.147	28.646	47.297	29.090	20.304
30	26.208	26.469	27.753	30.097	30.002	40.995	29.995	29.950	30.000	28.255	28.146
21	19.340	18.712	21.390	20.751	15.544	31.587	21.006	20.984	56.105	22.375	17.355
67	65.145	66.801	65.292	67.015	66.999	70.707	67.003	67.059	67.000	67.249	109.164
19	14.369	14.824	17.637	18.959	19.001	19.130	19.001	19.036	19.000	-37.528	10.937
60	54.662	55.451	54.364	59.848	60.001	65.389	54.438	37.725	60.000	57.332	84.021
19	25.617	25.803	23.755	19.191	21.324	28.758	19.030	19.127	19.000	14.075	19.111
90	88.385	89.436	90.345	90.001	90.001	70.238	89.977	90.061	90.000	30.053	127.978
27	29.533	30.057	33.956	36.664	32.251	30.749	27.012	6.144	27.000	27.638	16.195
40	37.879	40.233	36.174	39.959	40.001	51.679	24.885	40.080	40.000	35.383	52.265
25	21.596	21.618	23.713	35.960	25.001	23.858	50.190	25.037	25.000	24.570	17.986
30	24.814	24.722	26.966	30.038	30.000	22.066	30.004	30.118	30.000	27.588	23.503
hidden neurons	1	5	10	20	30	40	50	60	70	80	90
RMSE	3.719	3.410	3.744	3.679	1.970	8.424	7.898	8.872	13.865	21.536	15.645
MAE	3.371	2.890	3.118	1.539	0.916	6.998	3.966	4.238	6.615	11.482	10.126
MAPE	12.031	11.641	10.899	6.183	3.657	23.501	12.236	11.156	19.561	34.086	23.205

Table 6-6: Performance of test sample Model I using (nftool andBayesian Algorithm)

Target Output	Model no										
	I-46	I-47	I-48	I-49	I-50	I-51	I-52	I-53	I-54	I-55	I-56
24	21.998	22.997	24.308	23.997	22.428	24.644	23.532	23.446	24.247	20.883	24.430
64	59.487	66.006	63.592	64.568	65.084	64.206	62.675	64.476	64.106	62.648	63.472
115	112.003	113.070	115.238	122.962	115.408	115.128	114.020	114.699	114.847	81.705	115.300
26	26.956	26.231	25.258	26.186	25.901	25.857	25.179	27.197	25.384	9.556	25.216
9	11.273	9.953	9.386	9.639	10.269	8.706	9.825	9.892	9.350	5.697	9.085
62	57.501	59.505	61.349	60.860	60.992	61.966	60.352	61.129	62.307	54.458	60.882
23	27.490	25.875	26.534	27.994	26.766	33.938	26.163	26.204	25.150	29.090	24.518
30	25.365	26.615	29.170	28.321	28.628	26.150	26.936	29.143	29.488	28.255	13.309
21	18.501	16.032	16.505	17.814	16.547	20.710	17.537	17.519	18.470	22.375	18.298
67	65.008	66.694	67.269	67.113	66.817	66.969	67.231	67.175	66.928	67.249	66.954
19	13.712	17.217	18.840	17.184	17.508	18.921	17.434	19.659	18.858	-37.528	16.152
60	54.372	56.285	57.974	55.221	55.207	59.405	51.776	57.507	59.115	57.332	48.144
19	24.800	23.558	20.108	21.449	23.154	19.438	22.881	21.312	31.688	14.075	20.904
90	89.389	89.251	89.674	89.407	89.365	89.881	88.971	89.218	89.556	30.053	89.389
27	29.313	27.910	27.672	27.376	27.898	26.999	37.567	44.904	27.188	27.638	27.040
40	36.977	39.274	40.054	38.758	39.528	39.775	39.108	38.824	39.625	35.383	31.259
25	21.032	26.471	25.576	25.596	25.312	24.942	25.314	24.977	25.039	24.570	25.098
30	24.145	27.264	29.338	29.532	25.167	29.952	29.644	16.375	29.917	27.588	30.003
hidden neurons	1	5	10	20	30	40	50	60	70	80	90
RMSE	3.871	2.469	1.520	2.770	2.469	2.746	3.631	5.512	3.110	21.536	5.373
MAE	3.519	2.044	0.969	1.822	1.822	1.007	2.379	2.832	1.216	11.482	2.795
MAPE	12.442	7.388	3.767	5.535	7.078	4.089	8.221	10.460	5.714	34.086	8.426

Table 6-7: Performance of test sample Model I using (nftool and Scaled Conjugate Gradient Algorithm

Target Output	Model no											
	I-58	I-59	I-60	I-61	I-62	I-63	I-64	I-65	I-66	I-67	I-68	I-69
24	33.096	12.537	24.233	25.095	24.195	25.007	10.533	16.304	2.009	-5.154	20.993	25.280
64	66.278	53.469	63.008	77.050	82.479	57.345	72.820	61.306	41.515	65.922	72.758	127.234
115	136.012	110.106	117.389	115.567	112.932	117.213	125.198	119.275	120.705	106.863	161.911	118.284
26	30.125	20.901	27.583	26.872	21.741	28.228	31.621	21.762	17.526	12.018	23.289	30.290
9	22.263	8.225	10.540	8.918	10.211	18.246	19.023	20.785	38.770	-14.778	9.990	37.232
62	76.235	63.679	63.711	60.525	61.483	67.082	62.676	61.657	67.828	47.470	51.187	-18.605
23	35.309	30.002	32.020	27.362	26.008	26.182	27.279	19.822	41.607	44.601	25.762	22.017
30	22.727	24.976	32.056	28.682	26.147	29.055	23.556	34.030	1.169	13.711	8.466	32.861
21	31.107	23.206	19.928	16.488	21.051	20.633	21.974	11.974	50.438	36.203	3.843	3.549
67	57.491	55.502	61.495	78.401	77.008	61.870	29.601	71.658	42.579	68.239	74.133	69.878
19	19.618	15.910	8.412	18.510	17.756	10.390	31.989	24.514	16.583	8.614	-21.701	18.608
60	45.820	57.022	58.318	57.694	57.882	58.345	57.527	48.253	37.809	57.266	49.567	57.471
19	35.356	25.321	25.285	20.831	28.027	38.196	21.130	27.613	19.651	6.359	83.835	-3.124
90	40.081	82.132	88.663	88.869	86.937	121.852	68.468	89.184	71.764	87.537	94.066	30.887
27	28.603	31.772	32.478	27.493	25.836	28.038	21.539	25.804	2.381	39.040	38.107	22.675
40	52.779	30.428	40.477	40.610	39.403	34.237	34.620	37.118	25.640	28.440	38.854	45.224
25	14.988	23.513	16.695	25.530	24.764	1.877	49.733	57.464	14.747	39.079	36.354	-4.814
30	22.016	24.738	11.134	29.802	38.703	28.417	8.815	29.173	46.499	34.721	31.358	72.067
hidden neurons	1	5	10	20	30	40	50	60	70	80	90	100
RMSE	16.008	6.547	6.425	4.453	6.074	11.144	14.371	9.657	19.257	14.214	22.936	31.920
MAE	12.037	5.640	4.396	2.573	3.878	7.160	10.766	6.443	16.932	12.025	14.821	20.594
MAPE	36.611	16.810	16.702	6.522	11.112	25.718	34.679	28.003	64.145	53.859	52.962	63.630

Table 6-8: Selected Models for Model I Using MATLAB (nftool)

Model No	MAE	RMSE	MAPE (%)	Training Algorithm
Model I-38	1.970	0.916	3.657	Levenberg
Model I-48	**0.969**	**1.520**	**3.767**	**Bayesian**
Model I-61	2.573	4.453	6.522	Scaled

6.3 Developing of Model I Using Regression Excel Tool

A Linear regression model is used to statistically estimate the relationships among independent and dependent variables. The model is developed using Microsoft Excel (data analysis toolpak) in order to predict the payment delay for interim payment using 114 historical cases. Ninety-six cases are used for development and 18 cases are used for testing the model. Linear regression analysis is used to predict the output of dependent variable Interim payment delay (Y) on the basis of the independent variables which are: Delay in the submission of payment evaluation (X1); Claim review (X2); Payment process by owner (X3); Variation orders occurrence (X4); Extension of time occurrence (X5); Contractor experience (X6); Consultant's experience (X7); Externally funded projects (X8); Locally Funded projects (X9); Work progress (X10); Payment amount (X11). The data for independent and dependent variable are obtained from section (4-2). The linear regression output (Model I-70) is as presented by Equation (6-1), this equation is a result of the outputs shown in Table (6-9).

$$Y = -48.56 + 1.005*X1 + 1.022*X2 + 1.185*X3 - 0.268*X4 - 0.180*X5 + 0.766*X6 + 0*X7 + 1.484*X8 + 0.002*X9 + 0*X10 \quad \textbf{(6-1)}$$

The goodness of fit for the Regression Model is shown in Table (6-10) based on 95% confidence intervals. The Multiple R value represents the correlation coefficient of determination with a value of 0.996 showing a strong linear relationship between the predicted output and targets. The calculated R-squared value of 99.2% is representing the goodness of fit for the above equation. This means that 99.2% of the variance in the dependent variable (payment delay) is explained by the model indicating a high predictive power for the model. The adjusted R squared is equal to 0.991 and it represents R squared value in term of the number of variables in the model. As shown both values are very close to each other. The computed standard error is equal to 3.453%, this tells that the average distance for the predicted points falls about 3.453% from the regression line, this result asses a high precision of the prediction for the model. The obtained P-value of 2.054×10^{-103}(considered extremely significant) and reflecting the probability for obtaining an R squared value of 99.2%.

International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2016): 79.57 | Impact Factor (2015): 6.391

Table 6-9: Regression Coefficients for Model I-70using equation (6-1)

Code	Variables	Coefficients	Standard Error	Lower 95%	Upper 95%
C	Constant	-48.56	2.791	-54.1	-43.0
X1	Delay in the submission of payment evaluation	1.005	0.036	0.93	1.08
X2	Claim review(days)	1.022	0.012	0.99	1.05
X3	Payment process by owner (days)	0.981	0.013	0.96	1.01
X4	Variation orders occurrence (binary)	1.185	0.709	-0.22	2.59
X5	Extension of time occurrence (binary)	-0.268	1.049	-2.35	1.81
X6	Contractor Experience (years)	-0.180	0.148	-0.47	0.11
X7	Consultant's experience (years)	0.766	0.350	0.07	1.46
X8	Externally funded project (binary)	0.000	0.000	0.00	0.00
X9	Locally funded project (binary)	1.484	1.327	-1.14	4.12
X10	Work progress (delay in days)	0.002	0.007	-0.01	0.02
X11	Payment amount (BD)	0.000	2.72E-06	-4.49E-06	6.29E-06

Table **6-10**: Goodness of Fit of Model I-70using equation (6-1)

Regression Statistics	Goodness of Fit >= 0.80
Multiple R	0.996
R Square	0.992
Adjusted R Square	0.991
Standard Error	3.453
P-Value	2.054E-103
Observations	114

Table (6-9) shows the variables coefficients, the standard error along with the lower and upper bound of the confidence interval of each variable. The table shows that seven out of twelve variables (58.3%) has a positive biased coefficient, three out of twelve (25%) has negative biased coefficients and two out of twelve (16.6%) has zero value coefficients. As shown in the table for a 95% confidence prediction interval, about 95% of the observations should fall within Coefficient ± 2*Standard error from the regression line, these ranges of confidence intervals are expected to comprise the right value of the coefficient for each variable of the regression model.

Table (6-11) shows the P-values of the regression model which examine the effect of each independent variable on the dependent variable and its significance to the model. For a 95% confidence level a P-value of more than or equal to 0.05 is considered not significant, while P-value of less than 0.05 is considered significant. The results show that seven variables are considered as not significant, four out of these seven are binary (0, 1) data variables which are: externally funded projects, locally funded project, Variation orders occurrence (binary), and Extension of time occurrence (binary). This indicates that the model considers all the binary (0, 1) input data as not significant. The other non-significance factors are: Work Progress (Delay in days), Payment Amount (BD), and contractor experience. This is because these three factors do not have a direct effect on the payment delay. The Model consideration of "Extension of time" and "Variation" as non-significant factors may be explained by their non-direct relation with payment delay. Because a delay in the approval of extension of time and variation orders by the consultant causes a delay in issuance of payment certificate and as a result a delay in MoF process which results eventually in a payment delay to contractor. On the other hand, the non-direct relation of work progress to payment delay is expected because the delay in progress of work of a project might cause: 1) a delay in the documentations required to support payment or 2) a delay in payment evaluation, or 3) cause the occurrence of extension of time. These three might lead to delay in payment evaluation and review, a delay in issuance of payment certificate, a delay in MoF process, and eventually a delay in payment to contractor.

Table 6-11: Significant & Non-Significant Variables Model I-70using equation (6-1)

Code	Variables	t Stat	P-value	Significance
C	Constant	-17.40	1.406E-32	Significant
X1	Delay in the submission of payment evaluation	27.78	6.399E-50	Significant
X2	Claim review(days)	83.94	2.288E-97	Significant
X3	Payment process by owner (days)	75.63	9.819E-93	Significant
X4	Variation orders occurrence (binary)	1.67	0.097	Not Significant
X5	Extension of time occurrence (binary)	-0.26	0.79	Not Significant
X6	Contractor experience (years)	-1.22	0.22	Not Significant
X7	Consultant's experience (years)	2.19	0.031	Significant
X8	Externally funded project (binary)	0	1.00	Not Significant
X9	Locally funded project (binary)	1.12	0.26	Not Significant
X10	Work progress (delay in days)	0.28	0.78	Not Significant
X11	Payment amount (BD)	0.33	0.74	Not Significant

Accordingly, the model is modified by eliminating the non-significant factors and including only the significant independent factors as per the result of Table (6-11), which are: Delay in the submission of payment evaluation (X1), Claim review (X2), Payment process by owner (X3) and Consultant's Experience (X7), as presented in Equation (6-2).

Y = -48.96 +0.997*X1 +1.023*X2 +0.978*X3 +0.820*X7 (6-2)

The goodness of fit for the Regression Model is shown in Table (6-12) based on 95% confidence intervals. The Multiple R value is equal to 0.996 showing a very good linear relationship between the predicted output and targets. The calculated R-squared value is equal to 99.1% which is indicating a great predictive power of the model. The adjusted R squared is equal to 0.991. The computed standard error is equal to 3.453%, this result asses a high precision of the prediction by the model. The obtained P-value of 3.699×10^{-112}(considered extremely significant) and reflecting the probability for obtaining an R squared value of 99.1%.

Table 6-12: Goodness of fit of Model I-70 using equation (6-2)

Regression Statistics	
Multiple R	0.996
R Square	0.991
Adjusted R Square	0.991
Standard Error	3.450
P-Value	3.699E-112
Observations	114

Table 6-13: Regression coefficients for Model I-70 using equation (6-2)

Code	Variables	Coefficients	Standard Error	Lower 95%	Upper 95%
C	Constant	-48.961	1.882	-52.691	-45.230
X1	Delay in the submission of payment evaluation (claim)	0.997	0.035	0.927	1.067
X2	Claim review(days)	1.023	0.011	1.001	1.046
X3	Payment process by owner (days)	0.978	0.013	0.953	1.003
X7	Consultant's experience (years)	0.820	0.252	0.321	1.319

Table (6-13) shows the variables codes, coefficients, the standard error along with the lower and upper bounds of the confidence interval of each variable. The table shows that four out of five variables (80%) have a positive biased coefficient, and one out of five (20%) have negative biased coefficients. As shown in the table for a 95% confidence prediction interval, about 95% of the observations should fall within coefficient value ± 2*Standard error from the regression line, these ranges of confidence intervals are expected to comprise the right value of the coefficient for each variable of the regression model.

The test sample cases (18 cases) are used to validate the regression models of Equation (6-1) and Equation (6-2) for comparison purpose. The test sample along with the estimated output and the performance errors are shown in Tables (6-14) and (6-15) for Equations (6-1) and (6-2), respectively. As shown in Table (6-14), the performance error values for Equation (6-1) are: RMSE of 15.507, MAE of 3.655, and MAPE% of 12.291%. While Table (6-15) shows the performance error values for Equation (6-2) as follows: RMSE of 3.928, MAE of 3.655 and MAPE of 11.944%. Figure (6-1) shows the targeted and estimated interim payment delays (days) of the regression model for Equation (6-2). It is noticed that there is a remarkable matching between the two curves for Equation (6-2) results, which is reflecting very good performance for the regression model. On the other hand, Equation (6-1) results have slightly less accurate results as shown in Table (6-14). Therefore, it is recommended to choose Equation (6-2) for the regression model (Model I-70).

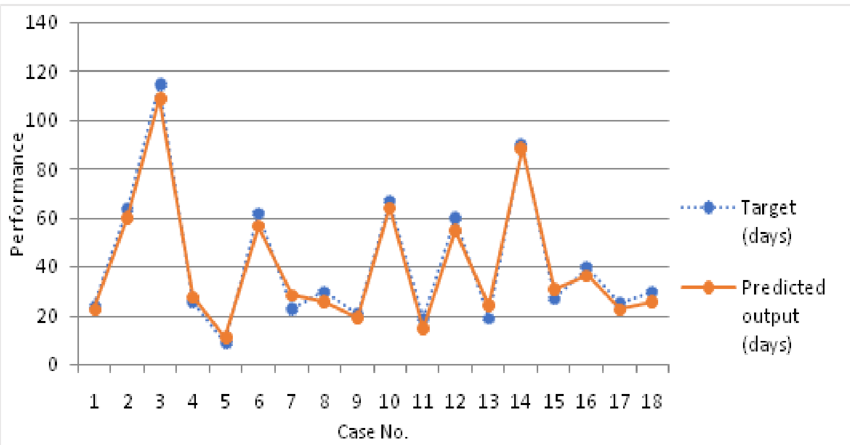

Figure 6-1: Performance of test sample of regression Model (Model I-70) using equation (6-2)

Table 6-14: Performance of test sample of regression model (Model I-70) using equation (6-1)

Case No.	Target output (days)	Predicted output (days)	Square Error	Absolute error	Percentage error%
1	24	22.303	0.400	0.094	0.004
2	64	59.595	1.038	0.245	0.004
3	115	109.260	1.353	0.319	0.003
4	26	27.601	0.377	0.089	0.003
5	9	11.058	0.485	0.114	0.013
6	62	57.986	0.946	0.223	0.004
7	23	28.291	1.247	0.294	0.013
8	30	25.844	0.979	0.231	0.008
9	21	19.074	0.454	0.107	0.005
10	67	64.933	0.487	0.115	0.002
11	19	14.359	1.094	0.258	0.014
12	60	54.905	1.201	0.283	0.005
13	19	25.457	1.522	0.359	0.019
14	90	87.311	0.634	0.149	0.002
15	27	30.790	0.893	0.211	0.008
16	40	37.489	0.592	0.139	0.003
17	25	22.044	0.697	0.164	0.007
18	30	25.302	1.107	0.261	0.009
Performance			RMSE	MAE	MAPE (%)
			15.507	3.655	12.291

Table 6-15: Performance of Test Sample of Regression Model (Model I-70) Using Equation (6-2)

Case No.	Target output (days)	Predicted output (days)	Square Error	Absolute error	Percentage error%
1	24	22.338	2.763	1.662	6.926
2	64	59.930	16.565	4.070	6.359
3	115	109.079	35.063	5.921	5.149
4	26	27.515	2.295	1.515	5.827
5	9	10.930	3.725	1.930	21.446
6	62	56.650	28.620	5.350	8.629
7	23	28.149	26.512	5.149	22.387
8	30	25.690	18.579	4.310	14.368
9	21	18.991	4.035	2.009	9.565
10	67	64.142	8.169	2.858	4.266
11	19	14.726	18.268	4.274	22.496
12	60	54.865	26.371	5.135	8.559
13	19	24.295	28.032	5.295	27.866
14	90	88.514	2.208	1.486	1.651
15	27	30.999	15.992	3.999	14.811
16	40	36.234	14.180	3.766	9.414
17	25	22.352	7.013	2.648	10.593
18	30	25.595	19.403	4.405	14.683
Performance			RMSE	MAE	MAPE (%)
			3.928	3.655	11.944

6.5 Comparison and Discussion of Model I Results

Table (6-16) shows the comparison between the selected models created for interim payment delay using MATLAB (nntool), MATLAB (nftool) and Excel Regression analysis tool. As per the computed results the best performance model is the neural network Model I-48 with 10 neurons in hidden layer based on Bayesian Regulation algorithm and tan-sigmoid function. Model I-48 has error values for RMSE of 1.520, MAE of 0.969, and MAPE of 3.767%.The ANN structure of the model chosen is shown in Figure (6-2). The best performance for the training sample of this model is at 120 epochs with MSE of 1.547 as shown in Figure (6-3), and R value of 0.999 which is reflecting a good fit of results and high predictive power of the network as shown in Figure (6-4). In the second place comes the regression model (Model I-70) with RMSE value of 3.777, MAE of 3.493, and MAPE value of 11.733%. Finally, in the third place the neural network Model I-22 with 100 neurons in hidden layer based on Bayesian Regulation algorithm and tan-sigmoid function. The ANN using MATLAB nftool shows relatively better results than the regression model. Although, on the basis of simplicity and ease of implementation of regression models, it will still give fairly good results for predicting of interim payment delay.

Table 6-16: Comparison of The selected models for (Model I)

Model No	MAE	RMSE	MAPE (%)	R-squared	Type	Tool
Model I-22	8.748	12.105	20.971	0.999	Neural Network	MATLAB nntool
Model I-48	0.969	1.520	3.767	0.999	Neural Network	MATLAB nftool
Model I-70	3.928	3.655	11.944	0.991	Regression Model	Excel ATP

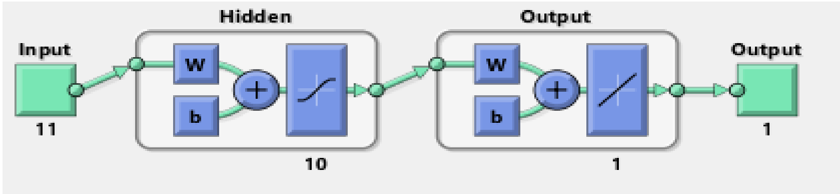

Figure 6-2: ANN Diagram for Model I-48

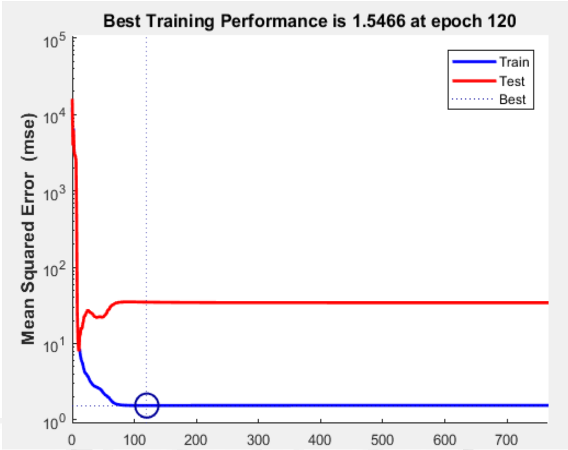

Figure 6-3: Training Performance R-Value Result for Model I-48

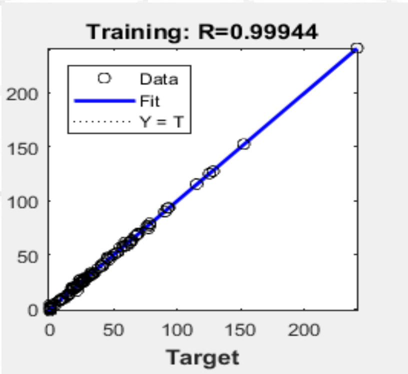

Figure 6-4: R square value for Model I-48

7. Development of Variation Payment Delay Model (Model V)

7.1 Introduction

The variation payment delay model (Model V) is developed for the prediction of variation payment delay using 91 actual historical cases for MoW building projects. The model is developed as neural network model using MATLAB nntool and MATLAB nftool approaches, and as a multiple linear regression model using Excel data analysis tool pack approach. A total of seventy-six (76) trial Models were created to come out with the best prediction model for variation payments delay.

7.2 Developing Model V Using MATLAB nntool

A total of thirty-six ANN models are trained, validated and tested for model V with different number of hidden neurons and training algorithms using MATLAB nntool. The ninety-one cases are divided into three sets: the training set consisting of 63 cases; the validation set consisting of 14 cases; and the testing set consisting of 14 cases. The performance of the test samples for multilayer (2 layers) feed-forward network with tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer for predicting interim payment delay are compared as shown in Tables (7-1), (7-2) and (7-3). The same training algorithms of Section 6-2 are used. By comparing the computed results for each of the three selected models using MATLAB (nntool), Model V-14 with 5 neurons in hidden layer based on Bayesian Regulation algorithm and tan-sigmoid function has been chosen. It has the lowest error values for interim payment prediction with RMSE of 0.304, MAE of 0.248 and MAPE value of 0.586% as shown in Table (7-4). The performance of the training and validation samples for the thirty-six trial models are not involved in the evaluation of models to eliminate the risk of misevaluation, because during the training and learning some models may accomplish good results with the training and validation data set, but very poor with new data, which is the test set. Therefore, the evaluation of the models' performance is done by comparing their predicted output over an independent test sample.

Table 7-1: Performance of test sample Model V using (nntool and Levenberg-Marquardt)

Target Output	Model no											
	V-1	V-2	V-3	V-4	V-5	V-6	V-7	V-8	V-9	V-10	V-11	V-12
83	79.23	82.99	82.42	80.16	85.58	79.40	73.98	87.43	79.21	89.01	67.77	80.57
70	64.41	70.98	68.25	54.39	69.67	64.71	66.20	70.01	70.14	72.78	56.57	69.51
30	31.83	29.24	33.27	26.87	29.89	24.31	41.53	27.72	28.21	30.73	31.19	30.01
25	28.66	25.33	30.68	24.02	26.46	21.92	38.39	24.93	25.78	27.23	28.67	26.91
21	24.58	19.32	24.20	19.49	21.32	26.88	25.04	21.00	22.14	21.56	20.57	21.30
32	31.27	32.14	30.83	24.66	37.79	25.77	33.89	32.02	85.72	32.48	28.11	31.48
142	154.53	142.57	155.40	208.99	142.14	135.63	134.20	142.00	142.33	146.54	131.75	141.96
63	59.51	62.74	58.69	57.75	63.54	47.86	68.31	63.07	56.96	62.22	57.14	62.13
83	81.40	84.14	79.11	90.49	82.81	67.71	85.47	82.95	82.26	86.10	74.04	83.39
36	35.36	35.60	36.79	32.88	34.78	28.00	46.21	35.61	36.55	36.30	33.74	36.52
83	81.39	82.49	85.61	88.51	86.47	81.20	82.65	96.71	83.77	85.41	74.00	81.83
68	64.15	64.91	69.64	64.65	64.82	61.05	63.28	68.00	68.24	69.84	59.72	68.00
69	65.13	66.62	71.20	65.18	67.11	62.34	65.19	70.71	68.02	71.35	61.53	67.89
19	24.51	20.24	35.35	19.06	21.56	19.16	20.98	18.99	39.12	31.52	28.66	35.00
hidden neurons	1	5	10	20	30	40	50	60	70	80	90	100
RMSE	4.709	1.292	6.288	18.816	2.34	7.652	6.909	3.926	15.466	4.243	8.308	4.392
MAE	3.732	0.964	4.346	9.072	1.698	6.439	5.739	1.625	6.508	2.9	7.114	1.84
MAPE	8.062	2.295	12.405	11.49	4.217	13.05	14.432	2.402	21.997	7.9	13.477	7.492

Table 7-2: Performance of test sample Model V using (nntool and Bayesian Algorithm)

Target Output	Model no											
	V13	**V14**	V15	V16	V17	V18	V19	V20	V21	V22	V23	V24
83	79.14	82.78	93.90	83.00	83.00	82.86	82.80	82.74	78.20	83.04	77.24	82.66
70	64.47	70.37	79.40	77.60	72.78	70.15	70.22	70.31	69.99	72.57	66.76	70.43
30	32.81	30.41	41.21	30.12	30.25	30.12	29.09	29.28	30.10	29.92	35.84	29.56
25	29.77	25.45	38.60	25.22	24.93	24.83	25.56	25.34	25.44	25.18	33.18	25.20
21	25.11	21.18	20.69	18.12	20.96	21.05	21.08	24.62	20.91	21.10	22.41	21.06
32	32.19	31.82	19.32	32.05	41.07	31.99	31.78	31.96	31.88	31.02	31.01	31.97
142	155.82	142.04	157.61	142.01	142.01	141.99	141.90	163.94	141.99	155.4	143.57	141.9
63	59.93	62.63	66.47	63.08	63.14	63.00	62.58	62.99	62.70	62.98	59.15	63.19
83	81.53	83.06	82.04	81.54	82.89	78.80	83.30	84.34	83.26	82.91	77.26	81.63
36	36.32	35.91	41.85	35.56	36.00	35.56	36.44	35.81	35.80	36.69	39.01	36.07
83	80.21	82.64	53.29	82.79	83.01	83.04	83.48	82.82	83.17	83.01	84.62	82.81
68	63.15	68.63	44.38	68.04	68.26	68.00	67.55	68.40	67.95	68.21	66.00	68.19
69	64.18	68.97	45.16	69.21	68.74	68.97	68.72	68.87	68.82	68.84	66.31	69.01
19	25.09	19.10	28.98	18.89	19.02	18.89	19.61	23.93	19.21	18.84	26.73	19.03
No of hidden neuron	1	**5**	10	20	30	40	50	60	70	80	90	100
RMSE	5.260	**0.304**	14.828	2.213	2.539	1.132	0.433	6.101	1.298	3.659	4.475	0.424
MAE	4.178	**0.248**	12.225	0.960	0.930	0.390	0.372	2.457	0.498	1.335	3.829	0.255
MAPE	9.138	**0.586**	25.067	2.165	2.491	0.620	0.999	4.747	0.839	1.512	10.392	0.464

Table 7-3: Performance of test sample Model V using (nntool and Scaled Conjugate Gradient Algorithm)

Target Output	Model no											
	V25	V26	**V27**	V28	V29	V30	V31	V32	V33	V34	V35	V36
83	77.271	93.902	84.097	82.85	93.087	81.469	80.542	68.294	82.159	87.273	68.39	63.807
70	63.529	79.404	71.313	72.801	79.377	73.986	69.256	57.219	67.455	76.851	59.808	54.192
30	36.318	41.211	34.296	34.866	34.153	39.395	34.566	36.067	38.293	29.819	40.222	46.096
25	33.693	38.596	31.604	32.752	31.149	37.145	32.27	33.034	35.071	27.315	37.421	42.819
21	30.025	20.688	15.4	17.094	41.732	18.354	31.522	28.809	32.454	22.225	27.242	16.415
32	37.283	19.315	20.766	24.054	39.342	19.591	21.585	44.873	45.904	38.47	43.084	25.37
142	155.89	157.61	132.81	150.89	143.19	145.64	137.12	143.4	146.71	168.82	149.84	128.42
63	60.164	66.466	57.98	51.917	55.924	60.915	55.969	61.163	61.324	52.357	69.733	75.911
83	80.739	82.037	70.353	67.155	78.405	76.718	73.711	84.675	82.709	74.528	88.949	87.668
36	39.237	41.854	33.995	33.562	38.439	38.983	32.48	40.482	40.623	33.032	48.125	46.24
83	71.98	53.288	88.054	73.427	78.533	79.387	73.822	87.817	73.748	90.965	63.734	82.72
68	57.249	44.384	67.493	62.116	63.006	61.943	58.562	70.013	61.753	68.579	52.564	69.679
69	58.034	45.161	69.785	64.197	64.243	63.326	59.182	71.11	62.271	70.851	54.714	71.17
19	27.135	28.975	25.709	35.587	25.422	25.73	26.669	27.381	33.412	24.222	29.185	39.292
No of hidden neurons	1	5	**10**	20	30	40	50	60	70	80	90	100
RMSE	8.192	14.828	**6.367**	8.675	8.088	6.591	7.559	7.736	8.133	8.952	11.785	12.413
MAE	7.472	12.225	**5.147**	7.324	6.699	5.656	6.914	6.357	6.790	6.132	11.185	10.425
MAPE	17.518	25.067	**13.160**	18.501	18.819	15.532	17.576	17.147	21.120	10.407	25.540	27.862

Table 7-4: Selected models for Model V using Matlab (nntool)

Model No	RMSE	MAE	MAPE (%)	Training Algorithm
Model V-2	1.292	0.964	2.295	Levenberg
Model V-14	**0.304**	**0.248**	**0.586**	**Bayesian**
Model V-27	6.367	5.147	13.160	Scaled

7.3 Developing Model V Using MATLAB nftool

A total of thirty-nine ANN models are trained, validated and tested for model V using MATLAB nftool. The performance of the test samples for multilayer (2 layers) feed-forward network with different number of hidden neurons, tan-sigmoid transfer function in the hidden layer and linear transfer function in the output layer for predicting variation payment delay is compared as shown in Tables (7-5), (7-6), and (7-7). The training algorithms used are Levenberg-Marquardt algorithm, Bayesian Regulation and Scaled Conjugate Gradient. By comparing the computed results for each of the selected models using MATLAB (nftool), Model V-60 structure with 80 neurons in the hidden layer, based on Bayesian Regulation algorithm and tan-sigmoid function has been chosen as it has the lowest error values for variation payment prediction with RMSE of 0.066, MAE of 0.018 and MAPE value of 0.058% as shown in Table (7-8).The best performance for the training sample of this model is at 378 epochs with MSE of 1.287×10^{-11}, and R value of 1 which is reflecting a perfect of results and very high predictive power of the network as shown in Figure (9-8). The performance of the training and validation samples for the thirty-nine trial models are not involved in the evaluation of models to eliminate the risk of misevaluation, because during the training and learning some models may accomplish good results with the training and validation data set, but very poor with new data, which is the test set. Therefore, the evaluation of the models' performance is done by comparing their predicted output over an independent test sample as mentioned in Section 5.3.

Table 7-5: Performance of test sample Model V using (nftool and Levenberg-Marquardt)

Predicted output	Model no												
	V-37	**V-38**	**V-39**	**V-40**	**V-41**	**V-42**	**V-43**	**V-44**	**V-45**	**V-46**	**V-47**	**V-48**	**V-49**
83.00	83.39	82.28	83.34	83.27	83.00	83.00	91.23	83.08	83.00	85.93	83.00	83.01	83.00
70.00	70.01	69.52	69.50	68.35	70.00	70.00	85.72	60.98	72.83	67.50	70.00	70.01	77.79
30.00	29.31	30.97	29.61	30.59	31.96	24.54	28.73	30.06	29.40	15.76	30.00	28.07	29.99
25.00	24.50	26.67	24.89	25.61	26.97	20.16	24.96	25.06	25.00	10.77	25.00	25.02	24.99
21.00	21.01	23.22	25.11	21.15	21.00	21.00	20.83	21.07	21.00	17.84	21.00	21.02	21.00
32.00	32.02	32.21	35.72	29.56	27.00	32.00	31.83	32.09	32.00	59.09	10.37	32.01	32.00
142.00	141.97	144.72	142.35	142.31	142.60	167.31	141.86	207.65	142.00	190.96	142.00	142.01	142.00
63.00	63.03	62.86	62.15	63.47	67.18	63.00	59.96	75.87	63.00	60.66	63.00	63.02	62.99
83.00	83.02	83.15	82.63	84.67	83.00	83.00	76.00	100.02	83.00	81.02	83.00	83.01	83.00
36.00	36.00	36.49	36.07	36.48	36.00	40.67	35.99	50.51	36.00	33.58	35.45	36.02	22.81
83.00	83.64	82.88	85.39	83.23	84.40	83.00	82.78	83.04	137.97	81.29	83.00	83.04	83.00
68.00	67.80	68.20	68.49	66.64	68.00	68.00	67.79	68.03	108.63	66.32	68.13	68.04	68.00
69.00	69.02	69.21	69.62	67.43	68.96	69.21	68.79	68.43	110.95	67.29	69.00	65.73	69.33
19.00	19.01	19.61	18.85	19.34	19.00	19.00	18.86	22.92	19.00	17.50	19.00	19.04	19.00
No of hidden neurons	1.00	5.00	10.00	15.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00
RMSE	0.31	1.12	1.66	1.11	1.94	7.15	5.18	19.04	21.45	16.01	5.78	1.02	4.10
MAE	0.18	0.78	1.03	0.87	1.08	2.89	2.61	8.86	10.07	9.03	1.59	0.39	1.53
MAPE	0.44	2.16	2.97	1.86	2.78	4.91	3.80	11.65	13.77	19.53	4.95	0.84	3.46

Table 7-6: Performance of test sample Model V using (nftool and Bayesian Algorithm)

Predicted output	Model No.												
	V-50	**V-51**	**V-52**	**V-53**	**V-54**	**V-55**	**V-56**	**V-57**	**V-58**	**V-59**	**V-60**	**V-61**	**V-62**
83.00	81.18	82.56	83.00	83.00	83.00	83.00	83.73	83.00	83.00	83.00	83.00	83.00	83.00
70.00	68.16	70.03	70.00	70.00	70.00	70.00	70.00	70.00	69.58	73.27	70.00	67.32	70.00
30.00	31.33	30.22	30.00	30.00	30.00	30.00	30.00	30.00	30.00	30.00	29.76	30.00	30.00
25.00	26.82	25.47	25.02	25.00	25.00	25.75	25.00	25.00	25.00	25.00	25.00	25.00	25.00
21.00	21.85	20.63	21.00	21.00	21.00	21.00	21.00	21.00	5.52	35.37	21.00	21.00	21.00
32.00	32.36	32.20	32.00	32.00	32.00	32.00	32.00	32.00	48.39	32.00	32.00	32.00	32.00
142.00	143.38	142.35	144.98	146.38	142.00	142.00	142.00	140.90	146.88	129.74	142.00	142.00	142.00
63.00	62.71	62.73	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	63.00	61.72
83.00	82.44	82.83	83.00	83.00	83.00	83.00	83.00	83.00	83.00	83.00	83.00	83.00	83.00
36.00	36.76	35.83	36.00	36.00	36.00	36.00	36.00	36.00	36.00	40.10	36.00	36.00	36.00
83.00	81.77	83.57	84.78	83.00	83.00	83.00	83.00	83.00	83.00	83.00	83.00	83.00	83.00
68.00	66.68	67.93	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00	68.00
69.00	67.68	69.03	69.40	69.00	69.00	69.00	69.00	69.00	69.00	69.00	69.00	69.00	69.00
19.00	20.51	18.96	19.00	18.08	19.00	19.00	19.00	17.47	19.00	19.00	19.00	19.00	19.00
hidden neurons	1.00	5.00	10.00	15.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00
RMSE	1.28	0.30	0.93	1.20	0.16	0.20	0.19	0.50	6.17	5.24	0.07	0.72	0.34
MAE	1.17	0.24	0.37	0.38	0.14	0.05	0.05	0.19	2.66	2.43	0.02	0.19	0.09
MAPE	2.80	0.57	0.35	0.57	0.30	0.21	0.06	0.63	9.21	6.65	0.06	0.27	0.15

Table 7-7: Performance of test sample Model V using (nftool and Scaled Conjugate Gradient Algorith

Predicted output	Model No.												
	V-63	**V-64**	**V-65**	**V-66**	**V-67**	**V-68**	**V-69**	**V-70**	**V-71**	**V-72**	**V-73**	**V-74**	**V-75**
83.00	82.28	84.04	62.05	79.35	78.37	81.95	82.57	64.16	91.30	69.95	79.19	119.60	85.31
70.00	69.52	73.08	51.16	69.83	63.89	69.52	67.71	54.01	93.72	52.95	70.09	94.75	68.84
30.00	30.97	33.67	35.35	31.02	32.79	26.96	28.03	26.29	30.58	25.03	23.86	22.03	31.08
25.00	26.67	29.48	30.55	26.66	27.86	22.54	23.83	24.20	30.22	20.05	19.63	16.82	27.47
21.00	23.22	15.50	32.80	20.65	18.23	21.12	21.36	21.23	22.04	94.47	23.82	1.87	20.01
32.00	32.21	29.56	29.44	40.82	36.24	59.58	72.42	29.30	41.79	21.78	19.78	42.33	33.26
142.00	144.72	118.07	132.38	141.50	151.69	148.01	142.76	141.76	143.73	144.85	138.69	135.63	143.93
63.00	62.86	63.54	71.40	60.84	64.34	54.89	60.47	63.30	53.64	46.10	69.26	59.89	63.54
83.00	83.15	80.78	88.03	80.49	84.50	77.76	82.50	72.01	59.56	86.81	89.40	82.35	91.00
36.00	36.49	37.89	45.27	35.70	36.43	36.79	37.26	52.57	39.24	39.89	39.40	43.77	29.55
83.00	82.88	84.72	80.68	79.52	79.80	83.46	82.87	88.07	83.94	68.40	79.15	82.21	81.43
68.00	68.20	70.67	68.02	65.73	69.67	71.18	67.21	65.64	71.26	52.22	75.51	69.80	68.65
69.00	69.21	71.81	69.11	66.01	70.31	72.65	68.65	65.94	68.09	53.16	66.06	69.98	72.75
19.00	19.61	18.86	18.45	22.00	41.55	33.86	18.62	-3.37	17.90	3.95	79.33	26.86	19.35
No of hidden neurons	1.00	5.00	10.00	15.00	20.00	30.00	40.00	50.00	60.00	70.00	80.00	90.00	100.00
RMSE	**1.12**	6.97	9.56	3.18	7.18	9.07	10.87	10.59	10.07	22.80	17.03	13.97	3.20
MAE	**0.78**	4.01	7.17	2.35	4.65	5.50	3.81	7.37	6.62	15.17	8.89	9.74	2.32
MAPE	**2.16**	7.58	15.30	5.66	14.31	15.93	11.11	18.73	12.04	46.10	33.23	24.31	4.75

Table 7-8: Selected models for Model V using Matlab (nftool)

Model No	RMSE	MAE	MAPE (%)	Training Algorithm
Model V-37	0.309	0.183	0.437	Levenberg
Model V-60	0.018	0.066	0.058	Bayesian
Model V-63	0.778	1.120	2.163	Scaled

7.4 Developing of Model V Using Regression Excel Tool

Linear regression analysis is used to predict the output of dependent variable (variation payment delay) on the basis of the independent variables, which are: Claim review (X1), Payment processed by employer (X2), Variations orders evaluation and approval (X3), Payment amount (X4), Balance of contingency in the contract (X5), Available contingency in the contract (X6), Extension of time occurrence (X7), Contractor experience (X8), Consultant experience (X9), and Work progress (X10). All the values of independent and dependent variables are known from historical data as explained previously in Section (4.2). Linear regression analysis is used to predict the output of dependent variable (Variation payments) on the basis of the independent variables mentioned above. The linear regression equation output is as presented in Equation (7-1):

$$Y = -68.825 + 0.92*X1 + 0.952*X2 + 0.984*X3 + 0*X4 + 0*X5 + 0*X6 - 1.193*X7 + 0.137*X8 + 0.095*X9 - 0.02*X10 \quad (7\text{-}1)$$

The goodness of fit for the Regression Model is shown in Table (7-9) based on 95% confidence interval. The Multiple R value represents the correlation coefficient of determination with a value of 0.999 showing a strong linear relationship between the predicted output and targets. The calculated R-squared value of 99.8% is representing the goodness of fit for the above equation. This means that 99.8% of the variance in the dependent variable (variation payment delay) is explained by the model indicating a high predictive power for the model. The adjusted R squared is equal to 99.7% and it represents R squared value in term of the number of variables in the model. As shown both values are very close to each other. The computed standard error is equal to 2.963%, this tells that the average distance for the predicted points falls about 2.963% from the regression line, this result shows very good prediction power of the model. The obtained P-value of 1.656×10^{-101} (considered extremely significant) and reflecting the probability for obtaining an R squared value of 99.8%.

Table (7-10) shows the variables codes, coefficients, the standard error along with the lower and upper bounds of the confidence interval of each variable. The table shows that five out of eleven variables (54.5%) have positive biased coefficients, three out of eleven (27%) have negative biased coefficients, and three have zero values. As shown in the table for a 95% confidence prediction interval, about 95% of the observations should fall within coefficient ± 2*standard error from the regression line, these ranges of confidence intervals are expected to comprise the right value of the coefficient for each variable of the regression model. Table (7-11) shows the P-values of the regression model which examine the effect of each independent variable on the dependent variable and its significance to the model. The results show that four variables are considered significant including the constant which are: Claim review, Payment by owner, and Variations orders evaluation and approval. These variables show very low p-values (less than 0.05) indicating a high significance. On the other hand, seven variables are considered as not significant showing high p-values (more than 0.05), one of these seven is extension of time, and it was input as binary (0, 1) variable and it is included in the regression equation with low coefficient amount. The most three non-significant variables are real variables entered in the model as payment amounts in Bahraini dinar which are: Payment amount, Balance of contingency in the contract, and Allocated contingency in the contract, these factors will not be included in the model because their coefficients are equal to zero. The other three non-significant factors are work progress, Contractor experience, and Consultant experience; these factors are included in the model with very low values of their coefficients. This may be due to their indirect relationship to payment delay.

Table 7-9: Goodness of Fit for Model V-76

Regression Statistics	Goodness of Fit >= 0.80
Multiple R	0.999
R Square	0.998
Adjusted R Square	0.997
Standard Error	2.963
P-value	1.656E-101
Observations	91

Table 7-10: Regression coefficients for Model V-76

Code	**Variables**	**Coefficient**	**Standard Error**	**Lower 95%**	**Upper 95%**
C	Constant	-68.825	7.249	-83.250	-54.400
X1	Claim review(days)	0.920	0.044	0.833	1.007
X2	Payment process by owner (days)	0.952	0.021	0.909	0.994
X3	Variation orders evaluation and approval (days)	0.984	0.006	0.972	0.996
X4	Payment amount (BD)	0.000	0.000	0.000	0.000
X5	Balance of contingency in the contract (BD)	0.000	0.000	0.000	0.000
X6	Allocated contingency in the contract (BD)	0.000	0.000	0.000	0.000
X7	Extension of time (binary)	-1.193	1.610	-4.397	2.011
X8	Contractor experience (years)	0.137	0.205	-0.271	0.544
X9	Consultant experience (years)	0.095	0.377	-0.655	0.845
X10	Work Progress (delay in days)	-0.020	0.017	-0.055	0.014

Table 7-11: Significant and non-significant variables Model V-76

Code	**Variables**	**t Stat**	**P-value**	**Significance**
C	Constant	-9.495	9.378E-15	Significant
X1	Claim review(days)	21.073	2.147E-34	Significant
X2	Payment process by owner (days)	44.931	1.593E-58	Significant
X3	Variation orders evaluation and approval (days)	164.84	4.527E-103	Significant
X4	Payment amount	-0.241	0.810	Not Significant
X5	Balance of contingency in the contract	0.823	0.413	Not Significant
X6	Allocated contingency in the contract	0.481	0.632	Not Significant
X7	Extension of time	-0.741	0.461	Not Significant
X8	Contractor experience (years)	0.669	0.506	Not Significant
X9	Consultant experience (years)	0.252	0.802	Not Significant
X10	Work progress (delay in days)	-1.177	0.243	Not Significant

Accordingly, the model is modified by eliminating the non-significant factors and including only the significant independent factors as per the model result which are: Claim review (days), Payment process by owner (days) and Variation orders evaluation and approval (days), as presented in Equation (7-2).

$$Y = -63.647 + 0.924*X1 + 0.940*X2 + 0.982*X3 \qquad (7\text{-}2)$$

The goodness of fit for the Regression Model for equation (7-2) is shown in Table (7-12) based on 95% confidence intervals. The Multiple R value is equal to 0.999 showing a very good linear relationship between the predicted output and targets. The calculated R-squared value is equal to 99.8% which is indicating a great predictive power of the model. The adjusted R squared is equal to 0.998. The computed standard error is equal to 2.923%, this result asses a high precision of the prediction by the model. The obtained P-value of 5.9×10^{-114}(considered extremely significant) and reflecting the probability for obtaining an R squared value of 99.8%. Table (7-13) shows the variables codes, coefficients, the standard error along with the lower and upper bounds of the confidence interval of each variable. The table shows that three out of four variables (75%) have a positive biased coefficient, and one out of four (25%) have negative biased coefficients. As shown in the table for a 95% confidence prediction interval, about 95% of the observations should fall within coefficient value ± 2*Standard error from the regression line, these ranges of confidence intervals are expected to comprise the right value of the coefficient for each variable of the regression model.

Table 7-12: Goodness of Fit for Model V-76Using Equation (9-2)

Regression Statistics	
Multiple R	0.999
R Square	0.998
Adjusted R Square	0.998
Standard Error	2.923
P-value	5.9E-114
Observations	91

Table 7-13: Regression Coefficients for Model V-76Using Equation (9-2)

Code	Variables	Coefficient	Standard Error	Lower 95%	Upper 95%
C	Constant	-63.647	1.276	-66.182	-61.111
X1	Claim review(days)	0.924	0.033	0.858	0.989
X2	Payment process by owner (days)	0.940	0.018	0.904	0.977
X3	Variation orders evaluation and approval (days)	0.982	0.005	0.972	0.993

The test sample cases (14 cases) are used to validate the regression models of Equation (7-1) and Equation (7-2). The test sample along with the estimated output and the performance errors are shown in Table (7-14) and (7-15) for Equation (7-1) and (7-2), respectively. As shown in Table (7-14), the performance error values for Equation (7-1) are for RMSE of 0.918, MAE of 0.842, and MAPE of 1.943. While Table (7-15) shows the performance error values for Equation (7-2) as follows: RMSE of 1.110, MAE of 0.958 and MAPE of 2.806%. It is noticed that there is a remarkable matching between the two curves for Equation (7-1), which is reflecting very good performance for the regression equation. On the other hand, Equation (7-2) results are showing relatively less matching. Therefore, it is recommended to choose Equation (7-1) for the regression model (Model V-76). Thus, Equation (7-1) is modified by excluding the variables with zero coefficient, it becomes:

$$Y = -68.825 + 0.92*X1 + 0.952*X2 +0.984* X3 -1.193*X7 + 0.137*X8 + 0.095*X9 - 0.02*X10 \qquad (7\text{-}3)$$

Table 7-14: Test Sample of The Regression Model (Model V-76) Using Equation (7-1)

Case No.	Target output (days)	Predicted output (days)	Square Error	Absolute error	Percentage error%
1	83	81.447	2.412	1.553	0.019
2	70	68.660	1.796	1.340	0.019
3	30	31.002	1.004	1.002	0.033
4	25	26.088	1.184	1.088	0.044
5	21	21.937	0.878	0.937	0.045
6	32	32.878	0.771	0.878	0.027
7	142	141.167	0.694	0.833	0.006
8	63	63.755	0.570	0.755	0.012
9	83	83.331	0.110	0.331	0.004
10	36	37.099	1.207	1.099	0.031
11	83	82.255	0.555	0.745	0.009
12	68	67.465	0.286	0.535	0.008
13	69	68.453	0.299	0.547	0.008
14	19	19.152	0.023	0.152	0.008
Performance			RMSE	MAE	MAPE (%)
			0.918	0.842	1.943

Table 7-15: Test Sample of the Regression Model (Model V-76) Using Equation (7-2)

Case No.	Target output (days)	Predicted output (days)	Square Error	Absolute Error	Percentage Error (%)
1	83	82.007	0.986	0.993	1.196
2	70	69.241	0.576	0.759	1.084
3	30	31.327	1.761	1.327	4.423
4	25	26.417	2.008	1.417	5.668
5	21	23.081	4.331	2.081	9.910
6	32	33.697	2.880	1.697	5.303
7	142	141.605	0.156	0.395	0.278
8	63	64.027	1.055	1.027	1.630
9	83	83.667	0.445	0.667	0.804
10	36	37.513	2.289	1.513	4.203
11	83	82.547	0.205	0.453	0.546
12	68	67.817	0.033	0.183	0.269
13	69	68.799	0.040	0.201	0.291
14	19	19.699	0.489	0.699	3.679
Performance			RMSE	MAE	MAPE (%)
			1.110	0.958	2.806

7.5 Comparison and Discussion of Model V Results

Table (7-16) show the comparison between the selected models developed for variation payment delay using MATLAB (nntool), MATLAB (nftool) and Excel Regression analysis tool. As per the computed results the best performance model is the neural network Model V-60 using MATLAB nftool with 80 neurons in hidden layer as shown in Figure (7-1) based onBayesian Regulation algorithm and tan-sigmoid function. The best performance for the training sample of this model is at 145 epochs with MSE of 1.50×10^{-5} as shown in Figure (7-2), and R value of 0.999 as shown in Figure (7-3).Model V-60 is having an error values for RMSE of 0.066, MAE of 0.018 and MAPE of 0.058%. In the second place comes the neural network Model V-14 with 5 neurons based onBayesian Regulation algorithm and tan-sigmoid function. Model V-14 has a minimum value for RMSE of 0.304, MAE of 0.248 and MAPE of 0.586%. Finally, in the third place the regression model (Model V-76) with RMSE value of 0.918, MAE value of 0.842 and MAPE value of 1.943%. The ANN using MATLAB nftool shows relatively better results than Matlab nntool and regression model.

Table 7-16: Comparison Of the Selected Models for (Model V)

Model No	RMSE	MAE	MAPE%	R	Type	Tool
Model V-14	0.304	0.248	0.586	0.999	Neural Network	MATLAB nntool
Model V-60	0.066	0.018	0.058	1	Neural Network	MATLAB nftool
Model V-76	0.918	0.842	1.943	0.998	Regression Model	ExcelATP

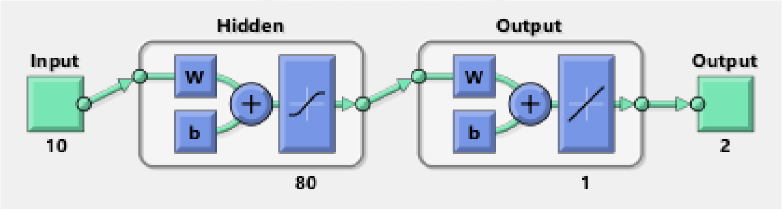

Figure 7-1: ANN Diagram for Model V-60

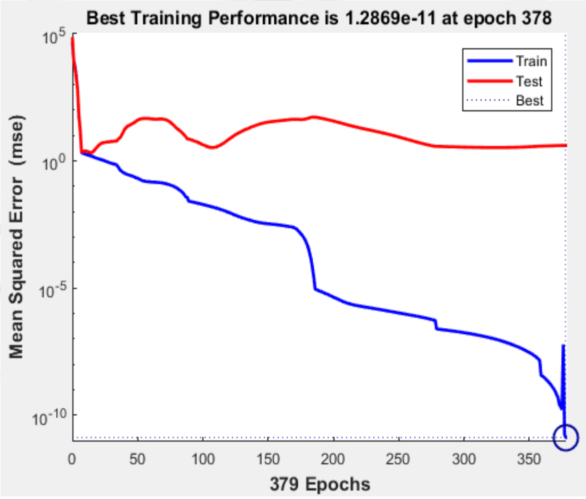

Figure 7-2: Training Performance For Model V-60

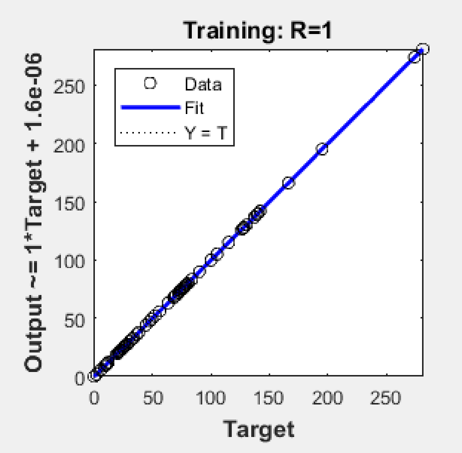

Figure 7-3: R square value for Model V-60

8. Conclusions and Recommendations

8.1 Conclusions

Two models are developed in this study for the prediction of interim payment delay (Model I) and prediction of variation payment delay (Model V). Models development is achieved through several steps starting by choosing the type of model, dividing the data set, choosing the software used to build the model, entering the dependent and independent variables, choosing the different parameters for each type of model and finally training and testing the models to choose the best model. Model I and Model V are developed using eleven and ten independent variables, for seventy and seventy-six trial models, respectively, using different parameters, programs and Modeling techniques including artificial neural models and regression models. For Model I for interim payment delay (days), the best performance model is a neural network Model I-48 developed using MATLAB nftool. This model is having 10 neurons in the hidden layer, and it is based onBayesian Regulation algorithm, and tan-sigmoid function. It has the minimum error results of RMSE of 1.520, MAE of 0.969 and MAPE of 3.767%. In the second place the regression model Model I-70 with RMSE of 3.928, MAE of 3.655 and MAPE of 11.944%. While for Model V for variation payment delay, the best performance model is neural network Model V-60 developed using MATLAB nfool. This model is having 80 neurons in the hidden layer and it is based onBayesian Regulation algorithm, and tan-sigmoid function. It has performance result of RMSE value of 0.066, MAE value of 0.018 and MAPE value of 0.058%. In the second place the regression model Model V-76 with RMSE value of 0.842, MAE value of 0.918 and MAPE value of 1.943%.

Finally, as the results show by comparing the Neural Networks and linear regression approach, it is shown that the estimation accuracy of Neural Networks approach gives relatively better results than linear regression analysis for payment delay (days) for governmental building construction projects.

8.2 Recommendations

This study has shown the important role each of the contract parties has on the occurrence of payment delay, some practices can increase the risk of payment delay while others can help in eliminating the risk, in this section some recommendations are presented to help contract parties in overcoming this issue:

1- MoW is recommended to use the developed models in this study to enhance the current payment process by predicting the payment delay risk in days and trying to avoid it by taking proper measures during the planning stage of projects.
2- It is recommended for MoW to use an electronic approval system to speed up the approval process and enhance communication. This can be implemented using an electronic signature to avoid the normal and slow cycle of hard documents approval.
3- Consultant is recommended to monitor the payments of the main contractor to the subcontractors and suppliers to avoid any payment delay to these parties in the project.
4- Contractor is recommended to use the models presented in this study in order to predict the delay in payment and to know when to make early claim of payments along with full required documents to avoid delay. Moreover, knowing the amount of delay risk for the contractor can help him to arrange the enough cash needed prior beginning of the project to avoid financial problems.
5- Contractor is recommended to use the model to have a good view about the expected payment delays in project, and to plan a cash flow programme that is applicable during construction stage.

8.3 Limitation of the Study:

The major limitation of this study is lack of data which is manifested in inadequate cases in payment delay. The study is also limited to one directorate of MoW, which is concerned with the building construction section; other directorates concerned with road and sanitary section are not covered. Another limitation is the scope of contractors being covered by the questionnaire by including only contractors of class A and B, and excluding other classes.

8.4 Future Recommended Studies

1- To Study the payment delay disputes of construction projects in Bahrain courts, and the remedies actions (verdicts) taken for those whose payments were delayed in each law case. A Model can be created thereafter to predict the remedies by ranking the factors taken from all these cases to create a clear picture for the payment delay compensation, and the most type of payment that is facing delay or nonpayment.
2- To develop payment delay models for governmental construction projects with more data and by including other types of payments such as advanced and final payment.

References

[1] Al-Sobiei, O., Arditi, D., and Polat, G. (2005). Predicting the risk of contractor default in Saudi Arabia utilizing artificial neural network (ANN) and genetic algorithm (GA) techniques. Construction Management and Economics, 23(4), 423-430.
[2] Amita, J., Singh, J., and Kumar, G. (2015). Prediction of Bus Travel Time Using Artificial Neural Network. International Journal for Traffic and Transport Engineering, 5(4), 410 - 424. http://dx.doi.org/10.7708/ijtte.2015.5(4).06.
[3] Ayman, A., A., H., Souma M., A., A., Ghaleb J., S., and Adnan, B. (2008), Prediction Model for Construction Cost and Duration in Jordan. Jordan Journal of Civil Engineering, 2(3), 250-266.
[4] Bordat, C., McCullouch, B., G., Labi, S., and Sinha, K., C. (2004) . "An analysis of cost overruns and time delays of INDOT projects." Tech. Rep. No. FHWA/IN/JTRP-04/7, Indiana Dept. of Transportation, In- dianapolis.[online] Docs.lib.purdue.edu. Available at: http:/docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1482&context=jtrp [Accessed 15 Jul. 2016].
[5] Börner K., Boyack, KW., Milojević, S., and Morris S. (2012). Basic model types, key definitions, and a general framework for the comparison of process models. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of science dynamics. Springer, Berlin, Ch 1, 3–22.
[6] Demuth, H., Beale, M., and Hagan, M. (2009). Neural network toolbox. 6th ed. Natick, Mass.: MathWorks 2-28.
[7] Elsawy, I., Razek, H., and Abdel, M. (2011). "A Neural Network Model for Construction Projects Site Overhead Cost Estimating in Egypt" IJCSI International Journal of Computer Science Issues;. 8(1) No.1 273–283.
[8] Goyal, S., and Goyal, G. (2012). "Time – delay simulated artificial neural network models for predicting shelf life of processed cheese". International Journal of Intelligent Systems and Applications. 4(5), 30-37.
[9] Iranmanesh, S., and Zarezadeh, M. (2008). Application of Artificial Neural Network to Forecast Actual Cost of a Project to Improve Earned Value Management System. World Academy of Science, Engineering and Technology International Journal of Educational And Pedagogical Sciences, 2(6), 658-661. Retrieved from http://scholar.waset.org/1999.10/13246
[10] Karanci, H. (2010). A Comparative Study Of Regression Analysis, Neural Networks And Case – Based Reasoning For Early Range Cost Estimation Of Mass Housing Projects. Citeseerx.ist.psu.edu. Retrieved 19 November 2017, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.633.7391&rep=rep1&type=pdf
[11] Matlab. (2017). En.wikipedia.org. Retrieved 15 August 2017, fromhttps://en.wikipedia.org/wiki/MATLAB
[12] Mensah, I., Nani, G. and Adjei-Kumi, T. (2016). Development of a Model for Estimating the Duration of Bridge Construction Projects in Ghana, International Journal of Construction Engineering and Management, 5(2), 55-64. doi: 10.5923/j.ijcem.20160502.03.
[13] Naik, M., and Radhika, V. (2015). Time and Cost Analysis for Highway Road Construction Project Using Artificial Neural Networks. KICEM Journal of Construction Engineering and Project Management, 5(1), 26-31.
[14] Odabaşi, E. (2009). Models for Estimating Construction Duration: An Application for Selected Buildings on the Metu Campus. Unpublished Master Thesis, METU, Turkey. Etd.lib.metu.edu.tr. Retrieved 19 November 2017, from https://etd.lib.metu.edu.tr/upload/12610696/index.pdf
[15] Parminder, K. (2016). Artificial Neural Networks As An Effective Project Management Tool. [online] Available at: http://data.conferenceworld.in/ICSTM5/P516-526.pdf [Accessed 20 Aug. 2017].
[16] Regression analysis. (2017). En.wikipedia.org. Retrieved 3 October 2017, from https://en.wikipedia.org/wiki/Regression_analysis
[17] Renuka, S., Kamal, S. and Umarani, C. (2017). A Model to Estimate the Time Overrun Risk in Construction Projects. Theoretical and Empirical Researches in Urban Management, 12(2), 64-76.
[18] Rose, S., Spinks, N., and Canhoto, A. (2014). Management Research: Applying the Principles (1st ed.). UK: Routledge.
[19] Waziri, B., Kadai, B., and Jibrin, A. (2014). Functional Duration Models for Highway Construction Projects in Nigeria. Arid Zone Journal Of Engineering, Technology And Environment, 20, 13-23.
[20] Yahia, H., Hosny, H., and Abdel Razik, M. (2011). Time Contingency Assessment in Construction Projects in Egypt using Artificial Neural Networks Model. IJCSI International Journal of Computer Science Issues, Vol. 8 (Issue 4 No. 2), 523 - 530