Diyala Journal of Engineering Sciences

Journal homepage: https://djes.info/index.php/djes

ISSN: 1999-8716 (Print); 2616-6909 (Online)

Forecasting Final Cost and Final Time of Highway Construction Projects in Iraq Using Regression Analysis

Dhuha M. Dawood[1*], Abbas M. Abd[2] and Raquim N. Zehawi[2]

[1]Department of Civil Engineering, University of Diyala, 32001 Diyala, Iraq
[2]Department of Highway and Airport Engineering, University of Diyala, 32001 Diyala, Iraq

ARTICLE INFO

Article history:
Received August 3, 2022
Revised October 19, 2022
Accepted December 5, 2022
Available online December 15, 2023

Keywords:
Risk management
Change order
Time and cost overruns
Regression analysis
Highway projects

ABSTRACT

The construction process is fraught with risks and uncertainties, which grow in proportion to the size and complexity of the project. The time and cost overrun of building projects and highways is a critical issue because of the great risks that result from it, which affect the project completion process and sometimes put it to a halt. In this paper, build a mathematical model using multiple regression analysis bases of historical data from highway projects to predict cost and time deviation in highway projects and used to control the cost and time of highway projects and thus reduce the risks of these projects. The study covers a group of highway projects in Diyala Governorate, where the records of (25) projects that were completed during (2012-2021) were used to develop and validate the regression model. Based on the literature review there are Seven factors that are considered the most influential on the final cost and final time. Hence, they are used as input parameters for the model. These factors are; bid cost, bid duration, change order cost, change order duration, number of change order item, start time (year) and finish time (year). It was found that the developed regression model has the ability to forecast the final cost (FC) and final time (FT) for highway projects as outputs with very high goodness of fit and coefficient determination (R^2) of (100%) and (99%) for FC and FT respective. it achieves significant savings in time and cost and to estimate and control the cost and time of highway construction projects.

1. Introduction

The construction process is fraught with risks and uncertainties, which grow in proportion to the size and complexity of the project. The project is considered successful if it is completed within the specified time and cost and with the required quality standards. Risk assessment consisting of risk identification, risk classification and risk analysis or evaluation is necessary for maintaining cost and quality of the project and for scheduled completion of the project [1] Therefore the advantages of insurance are increased and disaster losses to highway structures are decreased through risk analysis and evaluation [2]. Reference [3] considering most appropriate risk analysis methods employed in project risk management. The time and cost overruns of highways is a critical issue that affects the project completion process and sometimes stop it. The final cost of construction projects increases when unanticipated circumstances arise during the implementation phase. The majority of these occurrences are uncontrollable circumstances and resorting to change orders and the resulting cost of additional time that widen the difference between the contract's award price and the cost of final completion. Change orders have a negative impact on not just project performance

[*] Corresponding author.
E-mail address: eng.dhuha00@gmail.com
DOI: 10.24237/djes.2023.160412

in terms of time and cost, but also on quality, health, safety, and professional relationships [4], it's one of the most important problems in the construction industry [5]. change orders often involve additional cost and disruption to work already underway, leading to cost and time overruns, quality degradation, and loss in productivity on construction projects. Identifying factors that contribute in cost overrun can potentially lead to sustainable highway preservation strategies [6]. To that end, cost overrun has been widely studied in the literature [7, 8, 9, 10, 11, 12, and 13]. One of the most pressing issues in the construction sector of highways and buildings in countries appears to be risk management as a result of modification orders. There is limited research in this field; nevertheless, [14] discovered in their study that change orders are one of the most common causes of project delays, and that they represent a serious risk to project completion. This study makes an effort to use actual, quantifiable factors that are available before the project begins as predictors of the anticipated final cost as well as time and thus reduce the risks to which these projects are exposed. Change orders have been used, which is one of the ways to respond risks and to clarify the extent of their correlation and impact on the final cost and final time of highway projects, as the relationship between risk management and change order is one of the most pressing issues in the highway construction sector, and there is limited research very much in this field.

The goal of most parametric costing methods is to use some historical cost data and attempt to find a functional relationship between changes in cost and factors that caused to cost and time overrun and affect the final cost. The models of multiple regression estimates are well known and frequently applied in cost estimation. It works well because it follows clear mathematical methods and can explain the importance of each variable and how it is related to other variables. Regression models essentially find for the variables' linear structures that best correspond with the dependent variables. The following is how the general regression equation is expressed [15].

$$Y = A0 + A1I1 + A2I2 + \ldots.. AnIn \qquad (1)$$

where Y is the total estimated final result, A0 is a constant estimated by regression analysis, A1, A2…. An are coefficients also estimated by regression analysis, given the availability of some relevant data I1, I2…. In as measured distinguishable variables that may help in estimating Y [16].

Reference [17] Conducted research on the risks of excessive competition and only granting low-bid contracts in the Canadian public sector. It was discovered that the bidding procedure is a reliable predictor of the overall cost and possible cost growth. Over the years, the information gathered from (169) awarded road development projects in the West Bank of Palestine was evaluated (2004 - 2008). The analysis came to the conclusion that cost deviation affects all road projects in Palestine (100%) in some way [18]. The development of a regression model with a coefficient of determination (R) of (0.96). Therefore, the cost deviation affects road projects as shown in these studies and in different countries, and this is what is studied in this research paper in Iraq, especially road projects in Diyala Governorate.

2. Methodology

The data used in this research paper was collected from the Directorate of Highways in Iraq, the Directorate of Diyala Highways in particular for highway projects. The data included 25 completed projects, With regard to the construction phase and the consequent conditions in construction projects in general and highway projects in particular, which can lead to risks, the most important of which is mismanagement or inefficiency, and this will appear clearly in the additional cost or defect in work scheduling and the length of the implementation period, Two models were made for the cost and time of highway projects and based on the same data the date include dependent variables final cost (FC) and the final duration (FD) and the independent variables, change order cost (OC), bid duration (BD), bid cost (BC), change order duration (OD), final duration (FD), finish time(FT), start time (ST) and number of order items (NO.OI) [19] as show in Table 1 and the size of the investigated projects as shown in Table 2 to find out the impact of additional costs and durations resulting from a risk, and change orders were used to respond to this risk on the final cost and final time of the highway projects.Backtracking technology is adopted to analyze historical cost and time data in order to provide a robust model to aid in budgeting and cost estimation as well as time before work begins. Statistical Package for Social Sciences SPSS and MS Excel are used to develop a suitable model. In order to remove linear trend from the data, transformations by taking the natural logs of some of the variables reapplied. Then a simple linear model is developed using in each run, the natural log of the final project cost (FC) and the final projects time (FT). Detailed methodology used in this study is shown in Figure 1.

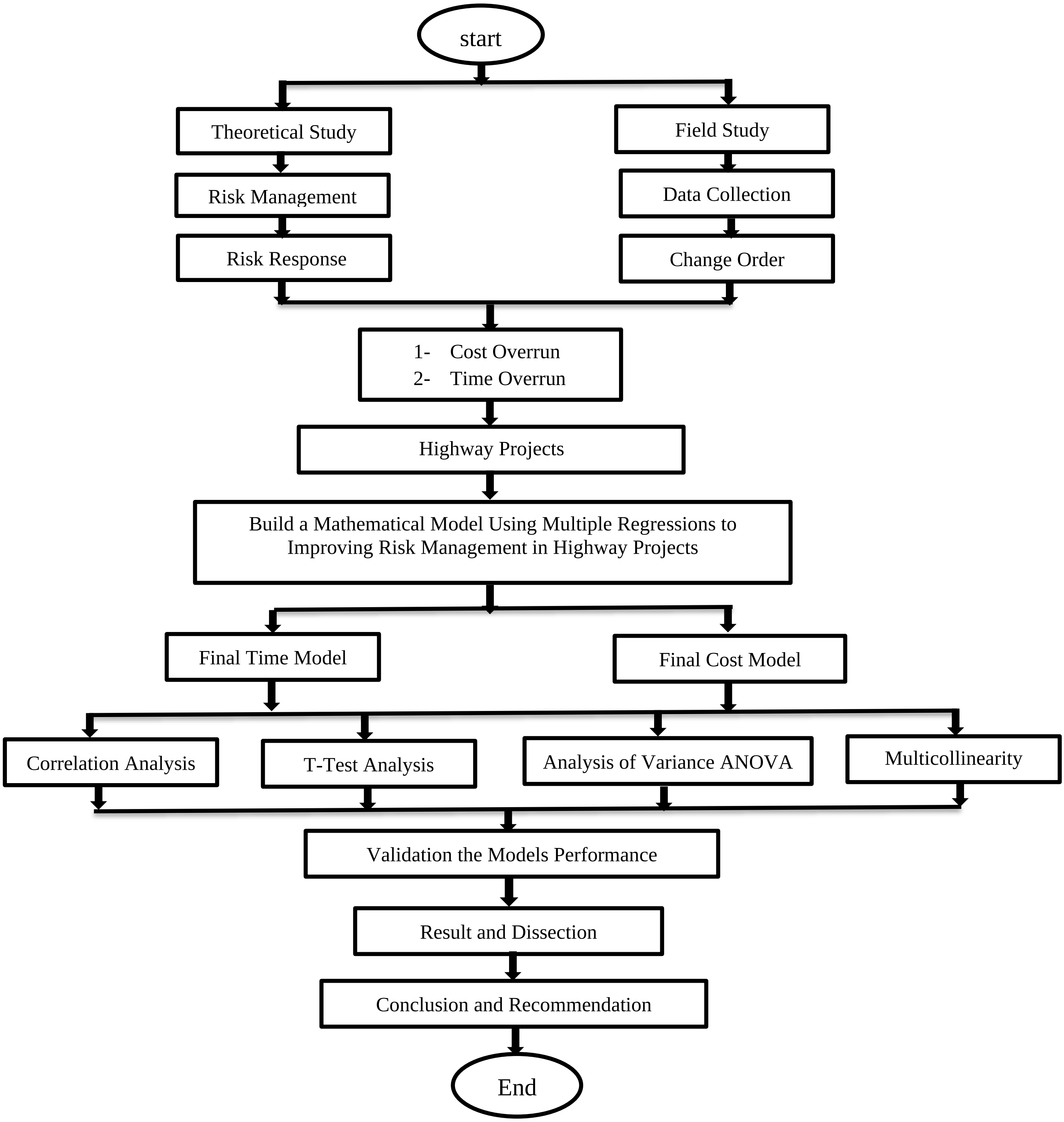

Figure 1. Schematic diagrams on the research design and procedure

Table 1: Collection of highway projects by researcher

No.	NP	BC	OC	BD	OD	ST	FT	NO.OI
1	Street maintenance and cladding	1010	76	232	20	15\9\2019	5\5\2020	3
2	Second lane road'	6386	954	365	124	03/12/2012	01/09/2020	6
3	road covering	1255	103	150	40	2\2\2014	13\5\2020	2
4	Paving a road	827	46	103	65	2\3\2020	13\6\2020	2
5	road rehabilitation	2581	209	365	135	28\9\2017	10\2\2019	5
6	road rehabilitation	8004	131	330	63	31\8\2017	27\9\2018	6
7	Road maintenance and rehabilitation	10851	654	365	130	18\4\2016	26\8\2017	5
8	Paving the road 3 km	345	33	90	20	11\11\2014	23\3\2015	2
9	Rehabilitation of village roads	1656	112	199	45	08/01/2020	25/07/2020	4
10	Rehabilitation and development of Baquba road	2790	265	210	40	01/09/2021	12/03/2021	3
11	paving villages roads	1465	145	210	70	17\7\2019	4\9\2020	4
12	road maintenance	1235	121	189	0	9\1\2020	16\7\2020	3
13	paving the road	299	28	90	50	8\11\2019	10\3\2020	3
14	Paving the road 311 km	1251	124	240	48	13\6\2019	28\11\2019	3
15	Maintenance and rehabilitation of the 20 km road	10597	235	365	125	18\4\2016	23\8\2017	4
16	paving the road	213	9	60	15	16\12\2020	2\2\2021	2
17	Complementing streets and roads	4669	17	240	90	19\1\2015	15\12\2015	5
18	Paving area with a length of 2 km	163	8	75	4	16\12\2020	5\3\2021	1
19	Street coverings 16 km	723	72	100	40	10\1\2021	30\5\2021	3
20	Paving and pouring waterwheels	545	54	120	30	5\11\2020	4\4\2021	4
21	Street rehabilitation	760	75	180	30	12\8\2021	30\11\2021	6
22	Construction of a 1000 m long road	129	12	150	30	25\5\2014	21\11\2014	1
23	Rehabilitation and covering of roads	173	17	90	0	21\1\2021	21\4\2021	2
24	paving the road	9109	1	360	90	19\11\2018	2\102019	6
25	paving the road	2267	368	320	110	7\11\2019	21\4\2021	7

NP Name of the projects

BC Contract cost in millions (ID)

OC	Total amounts change order in millions (ID)
BD	Contract duration (day)
OD	Total addition periods (day)
ST	The date of commencement Actual
FT	Contractual completion date, including available additional periods
NO. OI	Number of change order items

Table 2: The size of projects collected depending on the contract amount

Projects Size	150 - 1,000	1,000 - 5,000	5,000 - 10,000	10,000 - 15,000
NO. Of Projects	10	10	3	2
Note	The size of the projects was based on the contract amount in millions (1,000 This number is considered a [1 billion])			

3. Results and discussion

3.1 Correlation analysis

This analysis shows the correlation matrix between the independent variables and the dependent variable and between them and each other. When examining correlations between dependent variables and independent variables, the correlation coefficient's significance level should be equal to or lower than the P-value (0.01, 0.05), while when examining correlations between two independent variables, the correlation coefficient's value should be less than 0.7 in order to prevent lowering prediction accuracy [19].

Table (3) illustrates the value of the correlation coefficient and its significant level between the final cost (FC) and the independent variables, FC has a high correlation coefficient at a significant level less than (0.01) with; order cost (OC) its value (0.480), bid duration (BD) its value (0.886), bid cost (BC) its value (1.000), order duration (OD) its value (0.630), final duration (FD) its value (0.905), finish time(FT) its value (-0.166), start time (ST) its value (-0.324) and number of order items (NO.OI) its value (0.785). The positive relationship between OC, BC, NO. OI and the FC is a strong because it has a highest correlation than other variables.

Table 4 illustrates the value of the correlation coefficient and its significant level between the final time (FT) and the independent variables, FT has a high correlation coefficient at a significant level less than (0.01) with; order cost (OC) its value (0.516), bid duration (BD) its value (0.980), bid cost (BC) its value (0.901), order duration (OD) its value (0.712), final cost (FC) its value (0.905), start time (ST) its value (-0.400), finish time (FT) its value (-0.260) and no. of order items (NO.OI) its value (0.763). The positive relationship between OD, BD and the FD is a strong because it has a highest correlation than other variables were excluded due to their weak correlation for both models.

Table 3: Correlation analysis between FC and predictors

	FC	**OC**	**BC**	**BD**	**OD**	**FD**	**NO. OI**	**FT**	**ST**
FC	1.000	0.480	1.000	0.886	0.630	0.905	0.785	-0.166	-0.324
OC	0.480	1.000	0.463	0.500	0.342	0.516	0.435	0.108	-0.174
BC	1.000	0.463	1.000	0.882	0.629	0.901	0.779	-0.174	-0.323
BD	0.886	0.500	0.882	1.000	0.582	0.980	0.739	-0.243	-0.371
OD	0.630	0.342	0.629	0.582	1.000	0.712	0.557	-0.272	-0.416
FD	0.905	0.516	0.901	0.980	0.712	1.000	0.763	-0.260	-0.400
NO. OI	0.785	0.435	0.779	0.739	0.557	0.763	1.000	0.097	-0.050
FT	-0.166	0.108	-0.174	-0.243	-0.272	-0.260	0.097	1.000	0.706

ST	-0.324	-0.174	-0.323	-0.371	-0.416	-0.400	-0.050	0.706	1.000
FC		0.008	0.000	0.000	0.000	0.000	0.000	0.214	0.057
OC	0.008		0.010	0.005	0.047	0.004	0.015	0.304	0.202
BC	0.000	0.010		0.000	0.000	0.000	0.000	0.203	0.057
BD	0.000	0.005	0.000		0.001	0.000	0.000	0.121	0.034
OD	0.000	0.047	0.000	0.001		0.000	0.002	0.094	0.019
FD	0.000	0.004	0.000	0.000	0.000		0.000	0.105	0.024
NO. OI	0.000	0.015	0.000	0.000	0.002	0.000		0.323	0.406
FT	0.214	0.304	0.203	0.121	0.094	0.105	0.323		0.000
ST	0.057	0.202	0.057	0.034	0.019	0.024	0.406	0.000	

Table 4: Correlation analysis between FT and predictors

	FT	**OC**	**BD**	**BC**	**OD**	**FC**	**NO. OI**	**FT**	**ST**
FT	1.000	0.516	0.980	0.901	0.712	0.905	0.763	-0.260	-0.400
OC	0.516	1.000	0.500	0.463	0.342	0.480	0.435	0.108	-0.174
BD	0.980	0.500	1.000	0.882	0.582	0.886	0.739	-0.243	-0.371
BC	0.901	0.463	0.882	1.000	0.629	1.000	0.779	-0.174	-0.323
OD	0.712	0.342	0.582	0.629	1.000	0.630	0.557	-0.272	-0.416
FC	0.905	0.480	0.886	1.000	0.630	1.000	0.785	-0.166	-0.324
NO. OI	0.763	0.435	0.739	0.779	0.557	0.785	1.000	0.097	-0.050
FT	-0.260	0.108	-0.243	-0.174	-0.272	-0.166	0.097	1.000	0.706
ST	-0.400	-0.174	-0.371	-0.323	-0.416	-0.324	-0.050	0.706	1.000
FT		0.004	0.000	0.000	0.000	0.000	0.000	0.105	0.024
OC	0.004		0.005	0.010	0.047	0.008	0.015	0.304	0.202
BD	0.000	0.005		0.000	0.001	0.000	0.000	0.121	0.034
BC	0.000	0.010	0.000		0.000	0.000	0.000	0.203	0.057
OD	0.000	0.047	0.001	0.000		0.000	0.002	0.094	0.019
FC	0.000	0.008	0.000	0.000	0.000		0.000	0.214	0.057
NO. OI	0.000	0.015	0.000	0.000	0.002	0.000		0.323	0.406
FT	0.105	0.304	0.121	0.203	0.094	0.214	0.323		0.000
ST	0.024	0.202	0.034	0.057	0.019	0.057	0.406	0.000	

3.2 T-Test analysis

A stepwise technique was used to exclude variable that there is no high link shown the Tables 5 and 6 variables entered and removed for FC and FT models. T-test was used to statistically examine the coefficient of the generated model. Because of the significance of the t-test for the constant and the OC, NO. OI and BC coefficient being less than 1%, the summary of test table (7) illustrates a significant value to the predictor's coefficient for FC models and the OD and BD coefficient being less than 1%, the summary of test Table 8 illustrates a significant value to the predictor's coefficient for FT models.

Table 5: The FC Variables Entered/Removed

Model	Variables Entered	Variables Removed	Method
1	BC		Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	OC		Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	NO. OI		Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: FC

Table 6: The FT Variables Entered/Removed

Model	**Variables Entered**	**Variables Removed**	**Method**
1	BD		Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	OD		Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: FT

The Tables 5 and 6 illustrate the names of the variables that were included in the regression equation, where there were three variables in FC are OC, NO. OI and BC, and in FT there was two variables, OD and BD, as well as the gradual way to exclude the variables. The selected variables possess the highest correlation and influence on the models that have been built and shown in the Tables 7 and 8 where they explain the coefficients for each variable.

Table 7: The FC T-Test analysis

Model	**Unstandardized B**	**Coefficients Std. Error**	**Standardized Coefficients Beta**	**t**	**Sig.**
(Constant)	0.156	0.028		5.568	0.000
BC	0.972	0.005	0.980	178.801	0.000
OC	0.019	0.003	0.021	5.558	0.000
NO. OI	0.031	0.014	0.012	2.308	0.031

Table 8: The FT T-Test Analysis

Model	**Unstandardized B**	**Coefficients Std. Error**	**Standardized Coefficients Beta**	**t**	**Sig.**
(Constant)	0.456	0.120		3.806	0.001
BD	0.891	0.027	0.856	33.547	0.000
OD	0.095	0.011	0.213	8.367	0.000

3.3 Analysis of Variance ANOVA

A stepwise variance ANOVA was used to statistically examine the value of R2 was 100, and Adjusted R2 was 100. The F-test was 29615.106 at significant value of 0.00 for FC and the value of R2 was 0 .99, and Adjusted R2 was 0 .99. The F-test was 1151.928 at significant value of 0.00 shown the Table 9 for FC and Table 10 for FT, the significant value less than 1% and this mean that F- calculated greater than F tabular.

Table 9: Analysis of Variance ANOVA for FC

ANOVA[a]

Model	Sum of Squares	df	Mean Square	F	Sig.
Regression	43.462	3	14.487	29615.1	.000[d]
Residual	0.010	21	0.000		
Total	43.472	24			

Table 10: Analysis of Variance ANOVA for FT

ANOVA[a]

Model	**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Regression	8.352	2	4.176	1151.928	.000[d]
Residual	0.080	22	0.004		
Total	8.432	24			

3.4 Assumption of residuals normality

Satisfaction the normality of residuals is main assumption for developing regression models. The assumption histogram, Scatterplot and P-P plot presented that the distribution of the residuals is approximately normal as shown in Figures 2, 3 and 4 for FC and in Figures 5, 6 and 7 for FT.

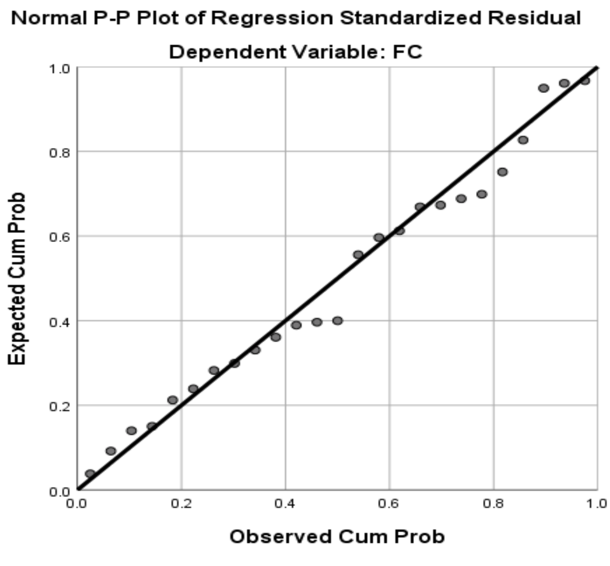

Figure 2. Normal P-P Plot of FC Regression

The P-P Plot in the Figure 2 illustrates that the data is gathered around the straight line and thus the bouquets are distributed by natural distribution and this is requirement of regression test conditions and the forecast of the prevalence of stations with the expected values and it is illustrated that there is no particular pattern of points and this is consistent with the requirement for regression test, this is illustrated in the Figure 3. In Figure 4 illustrates that the frequency runway fee is clear that the data follows the natural distribution and this is requirement of regression test.

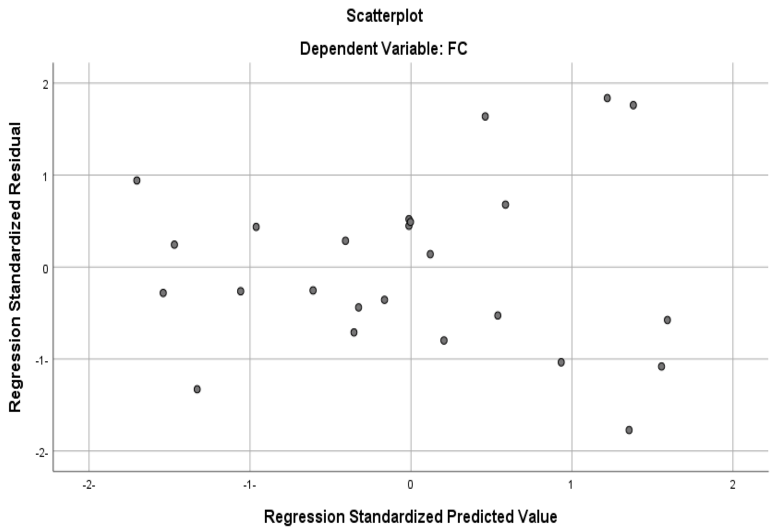

Figure 3. Scatterplot of FC Regression

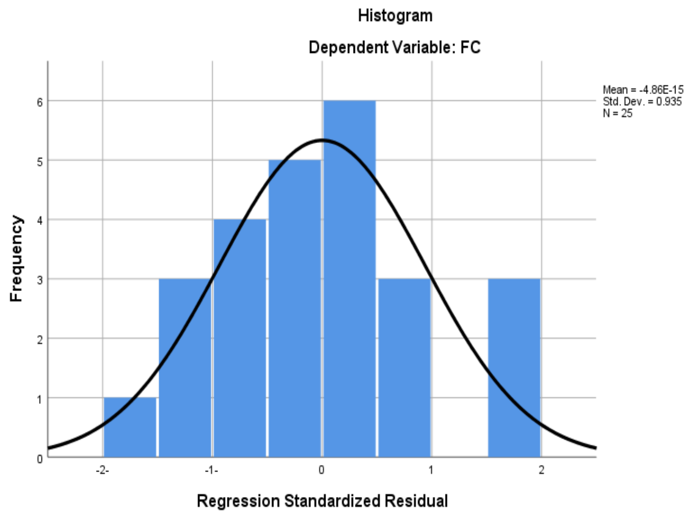

Figure 4. Histogram of FC Residuals

The P-P Plot in the Figure (5) for FT model also similar to FC model that shows the data is gathered around the straight line and thus the bouquets are distributed by natural distribution and this is requirement of regression test conditions. In Figure (6) the forecast of the prevalence of stations with the expected values and it is illustrated that there is no particular pattern of points and the frequency runway fee is clear that the data follows the natural distribution and this is requirement of regression test as shown in Figure (7).

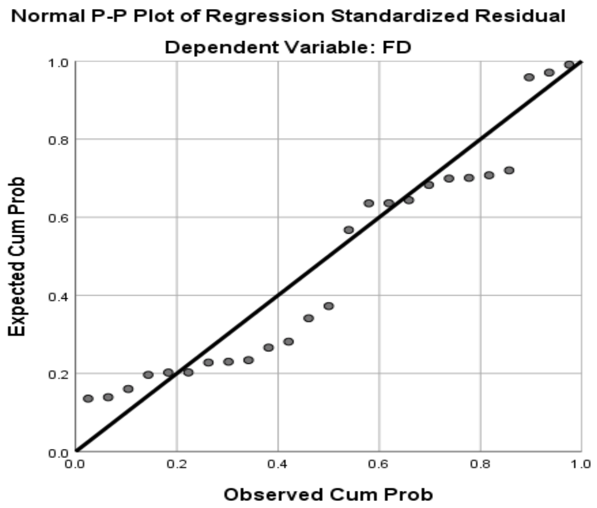

Figure 5. Normal P-P Plot of FT Regression

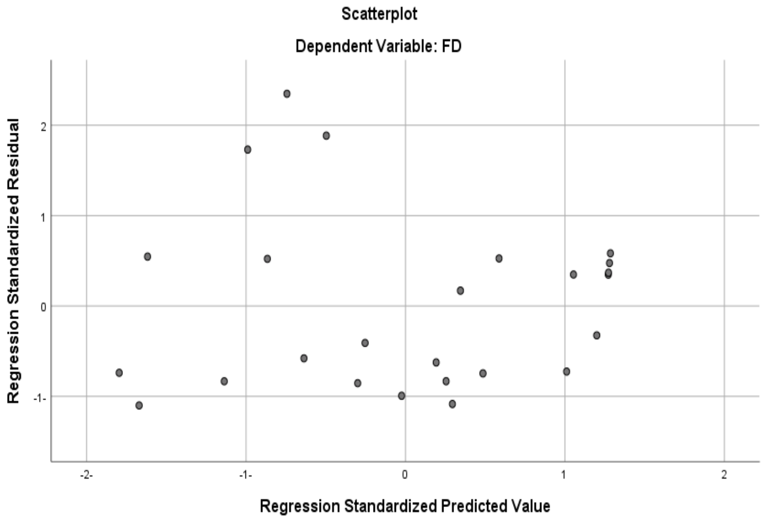

Figure 6. Scatterplot of FT Regression

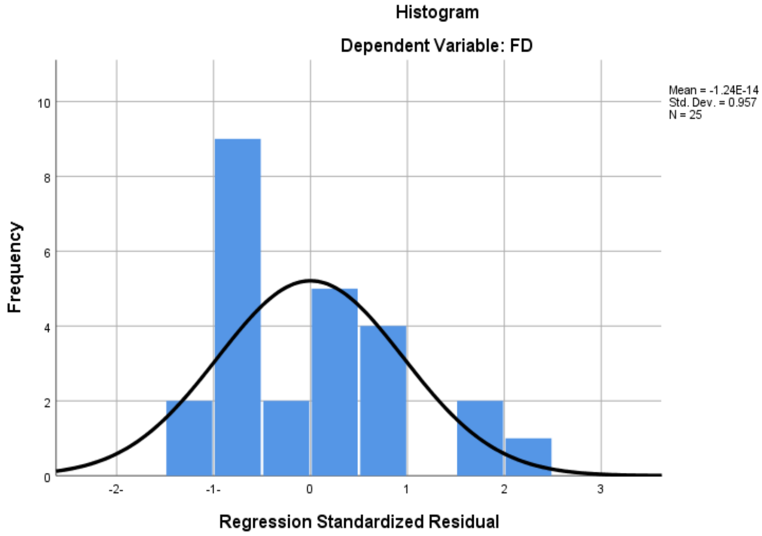

Figure 7. Histogram of FT Residuals

3.5 Multicollinearity analysis

In order to prevent erratic models, there must not be perfect correlation (i.e., no multicollinearity) between predictors. This hypothesis has already been investigated in the previously by tolerances and variance inflation factors (VIF) as shown in table (11) for FC and table (12) for FT. Multicollinearity is not a significant issue in this analysis because the forecasting of the FC and FT models must deal with (VIF) values that do not violate the specified criteria.

Table 11: The FC multicollinearity analysis

Model	**Collinearity Statistics Tolerance**	**VIF**
(Constant)		
BC	0.374	2.670
OC	0.772	1.296
NO. OI	0.386	2.589

From here it is clear that there is no problem of linear multiplicity between the variables, where the inflation coefficient was less than (3) and this also applies in the table (12) for FT model

Table 12: The FT Multicollinearity Analysis

Model	**Collinearity Statistics Tolerance**	**VIF**
(Constant)		
BD	0.661	1.513
OD	0.661	1.513

3.6 Validation the model's performance

In this paper, validation data are extracted from the same historical data file but for six additional randomly selected projects as shown in table (13). They are not part of the 25 projects used in developing the model. The model that developed by regression analysis have been assessed by Mean absolute error (MAE), Mean absolute percentage error (MAPE), and Root mean square error (RMSE) [20]. The data used for testing the model's performance include data of 6 highway construction projects for FC and FT models. Table (14) illustrates that most of validation criteria for FC model more than validation criteria for FT model; the highest difference was the RMSE for FC model which equal to 1.4502 more than the RMSE for FT model which equal to 0.4692.

Table 13: The additional six highway construction projects

No.	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
1	Paving the road 4 km	648	63	711	150	15	165	2021	2022	2
2	Paving the roads 3 km	423	1	424	120	15	135	2019	2019	2
3	paving village road 3200 km	286	28	314	90	19	109	2020	2020	2
4	Paving roads throughout the province	10,327	1,513	11840	540	538	1078	2019	2020	7
5	Paving the road 7 km	557	44	601	140	168	308	2011	2012	1

6	paving villages roads	557	54	611	80	25	105	2022	2022	3

2 Name of the projects
3 Contract cost in millions (ID)
4 Total amounts change order in millions (ID)
5 Final cost in millions (ID)
6 Contract duration (day)
7 Total addition periods (day)
8 Final duration (day)
9 The date of commencement Actual
10 Contractual completion date, including available additional periods
11 Number of change order items

Table 14: The Models Validation Criteria

Model	MAE	RMSE	MAPE
FC	0.745317753	1.450256526	0.113500066
FT	0.465325671	0.469205492	0.091134085

Through validation criteria in the above table, it is an application of MAE, MAPE, and RMSE. It is clear to us that these models have very high goodness of fit. It can be applied in highway projects because the amount of error in it is very small.

3.7 Regression Models for FC and FT

This section discusses the results of developing the regression methods models for FC and FT as follows:

1. The optimum method model built for FC.

$$\ln(FC) = 0.156 + 0.019 \ln(OC) + 0.972 \ln(BC) + 0.031 \ln(NO.OI) \quad (2)$$

This model is worth the attention as it appears goodness of fit because of its ability to predict 100% of the variation in the completed cost of the projects as the Adj.R2 value refers.

2. The optimum method model built for FT

$$\ln(FT) = 0.456 + .095 \ln(OD) + 0.891 \ln(BD) \quad (3)$$

This model can predict 99% of the variation in the completed duration of the projects as the high value of Adj.R2 refers. To produce typical statistics in standard form is to data were entered into a non-linear regression routine in the SPSS software. Similar to linear regression, the nonlinear function estimates the coefficients by minimizing the sum of the squared residuals. The difference between the regression types is that the nonlinear function requires initial parameter values to produce the desired statistics. As a result, the least squares coefficients presented in the above equation were entered into SPSS. It was found that the developed regression model has the ability to forecast the final cost (FC) and final time (F) for highway projects as outputs with very high goodness of fit and coefficient determination (R2) of (100%) and Adj.R2 (99%) for FC and coefficient determination (R2) of (100%) and Adj.R2 (91%) for FT respective.

Final equation models are:

$$FC = 1.168 * OC^{0.019} * BC^{0.972} * NO.OI^{0.031} \quad (4)$$

where
FC: final cost
OC: change order cost
BC: bid cost
NO. OI: number of change orders items

$$FD = 1.577 * OD^{0.095} * BD^{0.891} \quad (5)$$

FD: final duration
OD: change order duration
BD: bid duration

Table (15) and (16) illustrates the amount of prediction of FC and FT and comparing it with the actual. The prediction was much less than the actual, and this shows the amount of difference after using the models and the savings that these models will achieve when used in highway projects in terms of cost and time , in which it is clear that when using models, the cost and time of highway projects and the amount of change orders are controlled, thus reducing the increase in cost and time of highway projects that resulting from the process of responding to risks, thus improving the risk management process in these projects. The contract costs and time for highway projects can be calculated using these models with the best performance. The roads agencies in Iraq could use these models (Eq. (3) and (4)) above because these models have very high goodness of fit, it achieves significant savings in time and cost and to estimate and control the cost and time of highway construction projects and avoid the cost deviation after receiving change orders, which is one of the risk responses factors in projects because the two models have good accuracy and reliability.

Table 15: Comparison of predicted and actual FC

NO.	**BC. in millions (ID)**	**OC. in millions (ID)**	**BD. (day)**	**OD. (day)**	**FD. (day)**	**NO. OI**	**FC**	**P(FC)**
1	648	63	150	15	165	2	711	684
2	423	1	120	15	135	2	424	418
3	286	28	90	19	109	2	314	304
4	10,327	1,513	540	538	1078	7	11840	10723
5	557	44	140	168	308	1	601	587
6	557	54	80	25	105	3	611	589

Table 16: Comparison of predicted and actual FT

NO.	BC. in millions (ID)	OC. in millions (ID)	BD. (day)	OD. (day)	FD. (day)	NO. OI	FT	P(FT)
1	648	63	150	15	165	2	165	177
2	423	1	120	15	135	2	135	145
3	286	28	90	19	109	2	109	115
4	10,327	1,513	540	538	1078	7	1078	780
5	557	44	140	168	308	1	308	210
6	557	54	80	25	105	3	105	106

4. Conclusion

Through the analysis of the results the main conclusions indicate:

1. On the basis of historical data from highway projects, a regression model is developed using backward elimination techniques to determine the final highway cost perdition equation with

(R2 = 100%) and final highway time perdition equation with (R2 = 99%).

2. The most significant final highway cost model parameter is change order cost (OC), bid cost (BC) and final cost (NO. OI) and the most significant final highway duration model parameter are change order duration (OD) and final duration (BD).

3. The models developed in this study to forecast contract costs and time can be utilized for two different tasks. First, the contract costs and time for highway projects can be calculated using the model with the best performance.

4. The road agencies in Iraq should be used the FC and FT models for cost and time forecasting for highway construction projects to develop budgets and period for their projects plans and to control the cost and time than is being improved risk management in highway construction projects.

5. Highway agencies can create a user-friendly database to help with efficient data storing and extracting, as well as to prevent a poor job of data organization and storage, so they end up with a lot of important data they neglected it.

References

[1] Vishwakarma, A., Thakur, A., Singh, S., & Salunkhe, A. (2016). Risk assessment in construction of highway project. International Journal of Engineering Research & Technology, 5(2), 637-640.

[2] Ding, Y., Wang, P., Liu, X., Zhang, X., Hong, L., & Cao, Z. (2020). Risk assessment of highway structures in natural disaster for the property insurance. Natural Hazards, 104(3), 2663-2685.

[3] Kara Slimane, F., & Allal, M. A. (2020). Attempt to compare risk analysis methods in a highway viaduct project of the penetrating Ghazaouet in Algeria. Innovative Infrastructure Solutions, 5, 1-14.

[4] F. M. Arain and L. S. Pheng, "How design consultants perceive causes of variation orders for institutional buildings in Singapore," Architectural Engineering and Design Management, vol. 1, no. 3, pp. 181–196, 2005.

[5] Halwatura, R. U., & Ranasinghe, N. P. N. P. (2013). Causes of variation orders in road construction projects in Sri Lanka. International Scholarly Research Notices, 2013.

[6] Anastasopoulos, P. C., Haddock, J. E., & Peeta, S. (2014). Cost overrun in public-private partnerships: Toward sustainable highway maintenance and rehabilitation. Journal of Construction Engineering and Management, 140(6), 04014018.

[7] Al-Momani, A. H. (2000). "Construction delay: A quantitative analysis." Int. J. Proj. Manage., 18(1), 51–59

[8] Bhargava, A., Anastasopoulos, P. Ch., Labi, S., Sinha, K. C., and Mannering, F. L. (2010). "Three-stage least squares analysis of time and cost overruns in construction contracts." J. Constr. Eng. Manage., 10.1061/(ASCE)CO.1943-7862.0000225, 1207–1218.

[9] Chan, D. W. M., and Kumaraswamy, M. M. (1997). "A comparative study of causes of time overruns in Hong Kong construction projects." Int. J. Proj. Manage., 15(1), 55–63.

[10] Flyvbjerg, B., Shamris Holm, M. K., and Buhl, S. L. (2002). "Underestimating costs in public works projects: Error or lie?" J. Am. Plann. Assoc., 68(3), 279–295.

[11] Kara Slimane, F., & Allal, M. A. (2020). Attempt to compare risk analysis methods in a highway viaduct project of the penetrating Ghazaouet in Algeria. Innovative Infrastructure Solutions, 5, 1-14.

[12] Flyvbjerg, B., Shamris Holm, M. K., and Buhl, S. L. (2004). "What causes cost overrun in transport infrastructure projects?" Transp. Rev., 24(1), 3–18

[13] Kyte, C. A., Perfater, M. A., Haynes, S., and Lee, H. W. (2004). "Developing and Validating a Highway Construction Project Cost Estimation Tool." Final Rep., Virginia Transportation Research Council, Charlottesville, VA.

[14] Pathiranage, Y. L., & Halwatura, R. U. ENGINEER-Vol. XXXXI11, No. 04, pp.[17-30], 2010.© The Institution of Engineers, Sri Lanka.

[15] Thomas Bedford, "Analysis of the Low-Bid Award System in Public Sector Construction Procurement", MSc. thesis in Civil Engineering, 2009, University of Toronto.

[16] Stevens, I. James P., "Applied multivariate statistics for the social sciences" 2009, 5th edition-Taylor & Francis Group, LLC.

[17] Thomas Bedford, "Analysis of the Low Bid Award System in Public Sector Construction Procurement", MA. Thesis in Civil Engineering, 2009, University.

[18] Mahamid, Anna and Amund, B "Cost skew in road construction projects: the case of Palestine", Australian Journal of Building and Construction Economics, 12(1), 2012, pp. 58-71.

[19] Khaled, Z. S., Frayyeh, Q. J., & Aswed, G. K. (2015). Forecasting the final cost of iraqi public school projects using regression analysis. Engineering and Technology Journal, 33(2), 477-486.

[20] Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation Coefficients: Appropriate Use and Interpretation. Anesthesia & Analgesia, 126(5), 1763-1768.

[21] Chai, T., & Draxler, R. R. (2014). Root Mean Square Error (RMSE) Or Mean Absolute Error (MAE)? –Arguments Against Avoiding RMSE In the Literature. Geoscientific Model Development, 7(3), 1247-1250.