

Particle Swarm Optimization Based Approach for Estimation of Costs and Duration of Construction Projects

Tarq Zaed Khalaf ^a, Hakan Çağlar ^a, Arzu Çağlar ^b, Ammar N. Hanoon ^{c*}

^a Department of Engineering Management, Faculty of Engineering and Architecture, Kastamonu University, Turkey.

^b Sebahat-Mesut Yilmaz Vocational School, Kastamonu University, Turkey.

^c Department of Reconstruction and Projects, University of Baghdad, Baghdad, Iraq.

Received 26 October 2019; Accepted 12 January 2020

Abstract

Cost and duration estimation is essential for the success of construction projects. The importance of decision making in cost and duration estimation for building design processes points to a need for an estimation tool for both designers and project managers. Particle swarm optimization (PSO), as the tools of soft computing techniques, offer significant potential in this field. This study presents the proposal of an approach to the estimation of construction costs and duration of construction projects, which is based on PSO approach. The general applicability of PSO in the formulated problem with cost and duration estimation is examined. A series of 60 projects collected from constructed government projects were utilized to build the proposed models. Eight input parameters, such as volume of bricks, the volume of concrete, footing type, elevators number, total floors area, area of the ground floor, floors number, and security status are used in building the proposed model. The results displayed that the PSO models can be an alternative approach to evaluate the cost and-or duration of construction projects. The developed model provides high prediction accuracy, with a low mean (0.97 and 0.99) and CoV (10.87% and 4.94%) values. A comparison of the models' results indicated that predicting with PSO was importantly more precise.

Keywords: Cost; Duration; Construction Project; Particle Swarm Optimization; Managing Projects; Decision Making.

1. Introduction

The cost and duration prediction is considered an essential issue in construction projects. Underestimation and overestimation of costs may result from the failure of a construction project. The utilize of various approaches in the entire project lifetime should supply information on costs to the contributors to the project and support a complicated decision-making process [1, 2]. Cost and duration evaluating is a vital task for costing and tender preparation for any construction project before they are built. Cost and duration estimation in the early steps of construction projects comprises a considerable doubt. Hence, there is a high request to construct an active approach to minimize uncertainty in cost and duration prediction. One traditional technique for predicting the cost and duration values is the utilize of several specialists. Nevertheless, continuous contact with these specialists is not always an easy choice, producing to improve the alternative method to predict the cost and duration of construction projects. It is preferred to construct the new method according to datasets created from the preceding similar projects. Furthermore, utilizing the traditional technique is difficult and complicated. Hence, utilizing a soft-computing method is a noticeably, more effective way to address nonlinear problems. The best solutions for any system can be defined as the viable solutions with fitness values,

* Corresponding author: anh@uobaghdad.edu.iq



<http://dx.doi.org/10.28991/cej-2020-03091478>



© 2019 by the authors. Licensee C.E.J., Tehran, Iran. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).

as well as the values of any other sustainable solutions; these solutions are achieved by selecting values for the set of parameters that satisfy all constraint solutions [3]. Furthermore, optimization approaches are utilized widely in numerous areas, such as engineering and computer science. Study in the optimization area is very active, and new optimization approaches are being developed frequently [4]. The main goal of optimization methods is to find values for a set of parameters that maximize or minimize objective functions that are subject to certain constraints [5]. Through the last decades, numerous researches have been achieved to develop optimization approaches that apply evolutionary programming methods.

A relationship between completed construction cost and the time taken to complete a construction project was first mathematically established by Bromilow (1974) [6]. A regression analysis was utilized by Carr (1979) [7] to organize the duration and cost preparation of industrial buildings. Based on the neural network technique, Wang et al. (2013) [8] proposed a cost estimator model. The learning steps of their neural network were accomplished using a particle swarm optimization (PSO) method. In 2014, a hybrid model PSO-BPNN was proposed by Hong et al. (2014) [9] to assess the cost of construction projects. The PSO technique in the network has optimized the ANN weights. In 2015, Zima (2015) [10] presented a CBR model to predict the construction elements unit price. The CBR method shows a knowledge base that supports the cost prediction at the initial step of a construction project. A hybrid model ANN-ACO and ANN model for determining the amount and cost of construction waste in the early stage of construction were developed by Lee et al. (2016) [11] using "ant colony optimization" ACO algorithm to optimize the ANN weights and ANN model. In 2018, a proposed model for predicting the construction costs of sports areas was presented by Juszczak et al. (2018) [12]. Hybrid DES-PSO model that includes discrete event simulation (DES) and particle swarm optimization (PSO) algorithms were developed by Hegazy et al. (1994) [13] construction through a set of iterations in networks utilized, that significantly reduces efforts in search optimization scenarios.

The main target of the current research is to propose and investigate models of cost and duration estimating for the construction projects in the initial planning stage using particle swarm optimization (PSO) technique. The proposed PSO models can assist the engineers in making informed decisions in the initial stages of the design steps. With these models, it is probable to acquire a precise estimation, even when suitable information is not obtainable in the initial phases. These approaches encourage a feedback procedure that may support designers to attain the best solution. Moreover, the proposed models considered some category parameters, such as the security status that has been happened in Iraq in the last decade.

2. Research Methodology

Soft-computing methods are utilized to overcome complicated numerical optimization problems as non-linear systems. The current study tries to propose PSO models for predicting the cost and duration of construction projects accurately. The primary purpose of this study is to adopt and propose new models for the duration/cost assessment of construction projects utilizing the PSO algorithm. The proposed models were developed according to numerous, affecting input parameters, as presented below. The definition of the input parameters is listed in Table 1.

$$Cost/Duration = f(C, B, EN, FT, AGF, TFA, FN, SS) \quad (1)$$

Table 1. Definition of the input parameters

Parameter	Definition
<i>C</i>	The concrete volume: The concrete works comprise lean concrete, screed concrete, foundation, columns, beams, and slabs.
<i>B</i>	The brick volume.
<i>EN</i>	The number of elevators in the buildings.
<i>FT</i>	Types of footing: 1- Raft footing and 2-Separated footing.
<i>AGF</i>	The area of the ground floor.
<i>TFA</i>	The total area of floors.
<i>FN</i>	The floors number.
<i>SS</i>	The security status: 1- Safe, 2-Moderate and 3- Not safe

Optimization is required to produce optimal cost and-or duration values for a construction project. Of three key points must be taken in its progress:

- The objective function must be formulated.
- A clear approach is required to solve the optimization problem.
- The convergence criterion should be specified.

These detailed items will be discussed in the subsequent sections.

2.1. Objective Function of PSO Models

The primary objective of PSO is to optimize the cost and-or duration values and exploration for an optimum set of unknown coefficients, as illustrated in the proposed model section from within the solution space. The actual and forecast values of the duration and cost amount were detected to have minimal differences when using the final form of the optimized model. The proposed models are simulated utilizing MATLAB to optimize the duration and cost amount model for the construction projects. The objective function used in this study is the root mean square error (*RMSE*). This objective function can be accounted for utilizing the following expression [14-17]:

$$RMSE = \frac{1}{n} \sum_{i=1}^n |y - y'| \quad (2)$$

Where y' refers to the forecasted value, y refers to the actual value, and n denotes to the number of dataset samples.

2.2. Optimization Method of PSO Models

As a result of its global convergence ability, easy implementation, and adoption, PSO is considered one of the best optimization approaches. PSO is an evolutionary computation approach developed by Eberhart et al. (2001) [18], which was inspired by the social behavior of bird (particle) flocking. The PSO algorithm is generally accepted and used in solving different optimization problems. During the entire search process, the position and velocity of each particle can be updated according to Equations 3 and 4.

$$V_i(t+1) = wV_i(t) + c_1 \text{Rand}(\cdot)_1 |pbest_i t - X_i(t)| + c_2 \text{Rand}(\cdot)_2 |gbest_i t - X_i(t)| \quad (3)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (4)$$

Where V_i and X_i are the velocity and position of the particles, respectively; $\text{Rand}(\cdot)_1$ and $\text{Rand}(\cdot)_2$ are random numbers that are uniformly distributed between 0 and 1; $pbest$ denotes the best position of each particle in space, and $gbest$ represents the globally best position of all the particles. Acceleration coefficients c_1 and c_2 describe the 'trust' settings that mention the degree of confidence in the optimal solution found by an individual particle (c_1 -cognitive parameter) and by the whole swarm (c_2 -social parameter). The term w in Equation 3 refers to the inertial weight that was presented to improve the convergence of the iteration procedure. This weight is a scaling factor utilized to control the search capabilities of the swarm, which scales the current velocity value that affects the updated velocity vector. Later, Shi and Eberhart (1998) [19] developed the original PSO algorithm by adding the inertial weight; thus, this weight was not a portion of the original one.

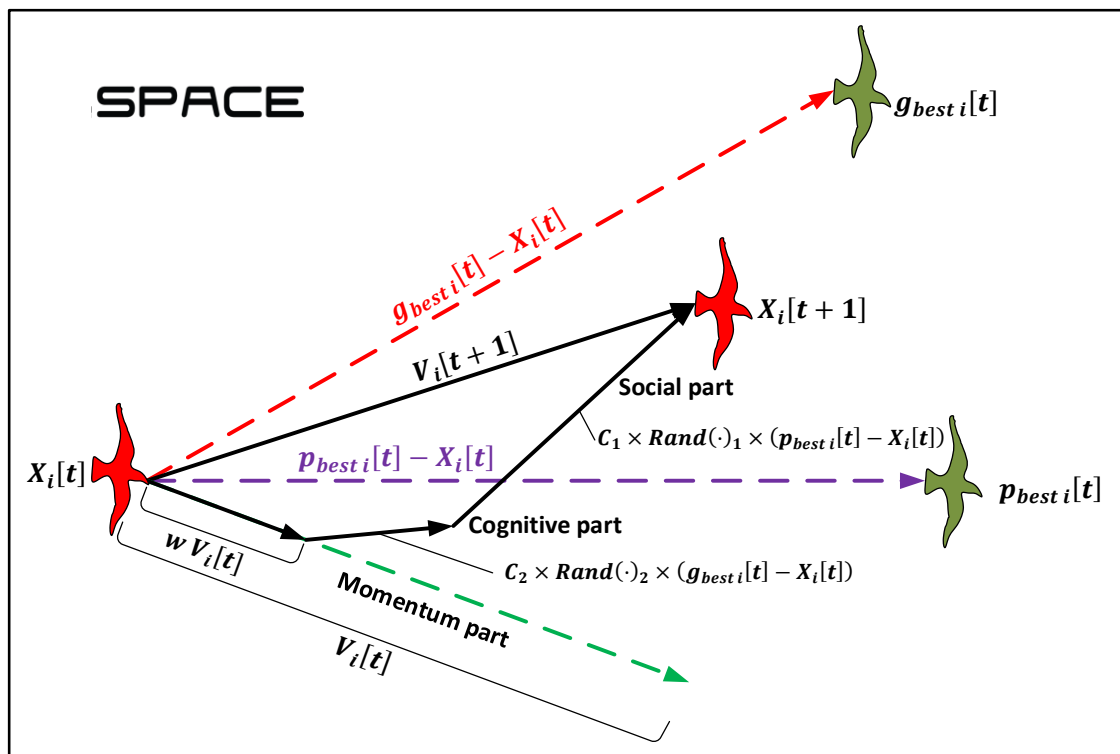


Figure 1. Pictorial view of particle behavior showing position and velocity update [20]

Figure 1 displays the updated particle position and velocity of the 2D parameter space. The first vector refers to the momentum velocity of the particles in the previous stage. The second vector refers to the particle memory components that get the best position as a result of iteration. This speed component attracted the particle to the best position in the solution space. So, the last vector is called a social component or swarm. The particle in this component is attracted to the best position in the swarm [20].

2.3. Convergence Criteria

Convergence criteria must be applied to end the process of optimization during the repeated search [21, 22]. The maximum number of iterations and minimum error requirements are the convergence criteria adopted in the PSO algorithm. The complexity of the optimization problem determines the maximum number of iterations. Previous knowledge of the optimal global error value determines the minimum error of the algorithm, which is possible to test or adjust the algorithm in mathematical problems when optimization is known a priori. Table 2 lists the main PSO parameters. Table 3 illustrates the convergence parameters of the PSO utilized in the current study [23].

Table 2. Main PSO parameters [23]

Description	Details
Number of particles, NP	A typical range is 10–40. For some difficult or special problems, the number can be increased to 50–100
The dimension of particles, n	It is determined by the problem to be optimized.
Inertia weight, w	Usually is set to a value less than 1, and for faster convergence, $w=0.7$ is considered.
Vectors containing the lower and upper bounds of the n design variables, respectively, x^L, x^U	They are determined by the problem to be optimized. Different ranges for different dimensions of particles can be applied in general
Cognitive and social parameters, c_1 and c_2	Usually $c_1 = c_2 = 1.494$. Other values can also be used, provided that $0 < c_1 + c_2 < 4$.

Table 3. PSO convergence parameters [23]

Description	Details
Maximum number of iterations (t_{max}) for the termination criterion	Determined by the complexity of the problem to be optimized, in conjunction with other PSO parameters (n, NP)
Number of iterations (k_f) for which the relative improvement of the objective function satisfies the convergence check	If the relative improvement of the objective function over the last k_f iterations (including the current iteration) is less or equal to fm , convergence has been achieved
Minimum relative improvement (fm) of the value of the objective function	

2.4. Proposed PSO Model

Figure 2 displays the flow chart of the proposed PSO model utilized in the current study. The correct choice of the PSO parameters has a remarkable effect on the performance of the algorithm. These parameters are namely: swarm size, coefficient of inertia, and acceleration coefficient. The neighborhood size is also needed for the local best algorithm.

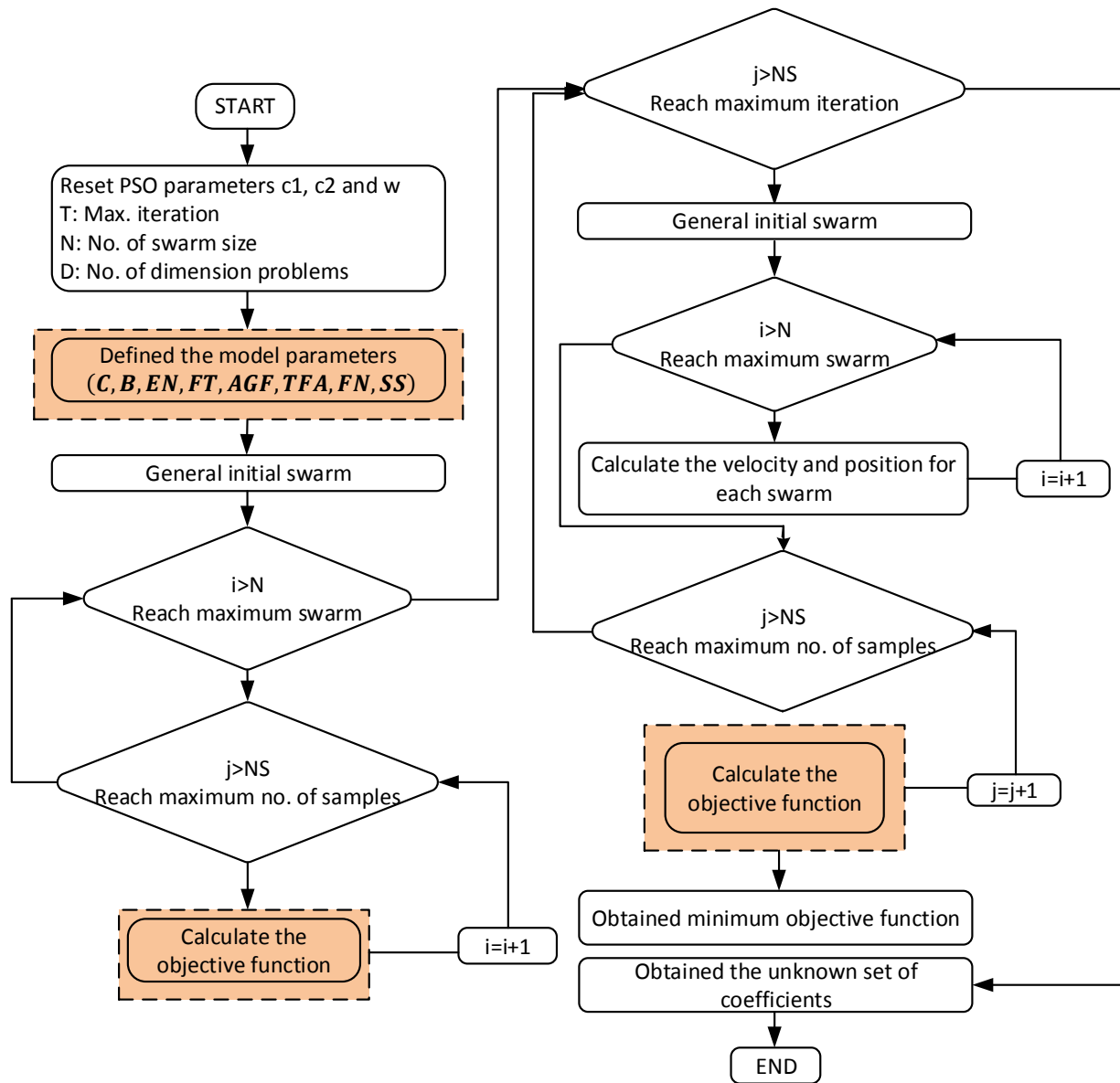


Figure 2. Flowchart of PSO of the proposed model

The optimization procedure typically uses a gradient-based algorithm appropriate for local exploration. Consequently, to be successful, the optimization procedure needs an initial point gotten from a global exploration. A strong training process requires both the initialization and optimization procedures. The following highlights how the PSO algorithm can be implemented to search for the optimum duration and cost amount of the construction projects.

- Create a swarm initialization by assigning a random location for each particle in the hyperspace problem.
- Evaluate the objective function of the proposed model for each particle.
- Compare the objective function value of each particle with $pbest$. If the current value is better than the $pbest$ value, this value is set as $pbest$, and the position of the current particle, X_i , is set to $pbest$.
- Identify particles with the best objective function value. The value of its target function is determined to be $gbest$, and its location is $gbest$.
- Update the velocity and the position of all particles based on Equations 3 and 4.
- Repeat steps 2-5 until the convergence criteria are met (the maximum number of iterations or a sufficient objective function value is obtained).

The proposed model was formulated utilizing MATLAB software to optimize the cost amount and duration models of construction projects. The proposed models to be optimized are as follows:

$$Cost = F_1 + F_2 \cdot SS + F_3 \cdot FN + F_4 \cdot C + F_5 \cdot B + F_6 \cdot EN + F_7 \cdot TA + F_8 \cdot AG + F_9 \cdot FT \quad (5)$$

$$Duration = K_1 + K_2 \cdot SS + K_3 \cdot FN + K_4 \cdot C + K_5 \cdot B + K_6 \cdot EN + K_7 \cdot TA + K_8 \cdot AG + K_9 \cdot FT \quad (6)$$

Where F_1 to F_9 and K_1 to K_9 are the unknown coefficients.

The main goal of utilizing PSO to optimize the cost amount and duration models is to examine for an optimum set of unknown coefficients. Hence, the difference between the actual cost amount of construction projects and that predicted utilizing the final form of the optimized expressions is minimal.

2.5. Description of Dataset

A total of 60 construction projects constructed by government contractors between 2008 and 2016 from different places in Iraq were collected. The selected projects (samples) represent about 80% of the projects implemented in Iraq in terms of implementation method, materials used, and architectural style. Eight input variables ($C, B, EN, FT, AGF, TFA, FN, SS$) and two output variables (cost amount or duration), as displayed in Table 4.

Models inferred using optimization tools have the capability to estimate within the data range obtainable and are applied for additional development. Thus, the size of the dataset utilized for the modeling process is essential, as it impacts the accuracy of the final models. The behavior of any model modified using this data is influenced by the sample size and its variable distributions. Therefore, the data is graphically illustrated in Figure 3 as histograms. Figure 3 depicts the statistics of the samples utilized in constructing the proposed model.

For high accuracy, the ratio of the number of dataset records to the number of input parameters should not be less than three, as proposed by Frank and Todeschini (1994) [24], and they recommended to be higher than five. For the present case study, this ratio was $60/8 = 7.5$, which exceed the recommended criteria. From the 60 samples (projects), 48 samples (80 %) were considered for building the proposed models, while 12 samples (20 %) were utilized to validate the proposed models. The descriptive statistics of the dataset utilized in this study are given in Table 5.

Table 4. Input and Output Parameters

Item	Project ID	SS	FN	C	B	EN	FT	AGF	GFA	Cost	Duration
1	Building of Hajj and Umrah/ Anah/Anbar	3	2	440	465	0	1	498	815	565.167	203
2	Event hall of the Ebad al-Rahman mosque/ Yusufiya / Baghdad	1	1	155	108	0	1	402	320	177.765	95
3	Secondary Sadiq Amin/ Muqdadiya/ Diyala	1	2	632	550	0	1	936	1610	844.313	204
4	Secondary Amjad Al Zahawi/ Baquba/ Diyala	1	2	563	412	0	1	716	1210	665.824	186
5	Secondary Prince Ali/ Tarmiyah/ suburb of Baghdad	3	2	612	638	0	1	833	1424	842.628	235
6	Secondary of the last prophets/ Yusufiya/Baghdad	1	2	481	508	0	1	661	1110	581.117	182
7	School of Habib Ben Khedi/ Nineveh	3	2	705	735	0	1	958	1650	1058.018	247
8	Abi Dardaa School/Kirkuk	1	2	815	850	0	1	1095	1900	1032.063	217
9	Falluja Islamic High School/Fallujah/Anbar	3	2	285	297	0	1	415	664	439.354	195
10	Abu Ghraib secondary school/Abu Ghraib/suburb of Baghdad	2	2	571	590	0	1	782	1330	811.354	211
11	Apartments for health staff/Ramadi/Anbar	3	1	312	220	0	1	765	650	495.48	134
12	Secondary Al-Moatasem/ Samarra/Salah al-Din	3	2	653	615	0	1	1095	1900	662.535	261
13	Residential Units for Employees/ Nukhayb / Anbar	3	2	98	140	0	1	239	344	237.018	177

14	Local Council Building/Wasit	1	2	429	444	1	1	556	920	636	174
15	Information Building for Local Council / Wasit	1	1	20.5	28	0	1	116	60	35.451	81
16	Halls for the pilgrims/Nukhayb/Anbar	3	1	256	235	0	1	545	450	324.805	122
17	Directorate of AL-Awqaf/Kut /Wasit	1	2	831	490	0	1	743	1260	768.365	189
18	Directorate of AL-Awqaf /Ishaqi/Salah Eddin	2	2	330	345	0	1	578	960	577.922	193
19	Directorate of AL-Awqaf/ Haditha/Anbar	3	2	345	340	0	1	578	960	660.144	210
20	College of Languages / University of Baghdad / Baghdad	2	2	6900	1550	0	1	4450	8000	5352.919	538
21	Expanding the building of the Faculty of Engineering /University of Baghdad /Baghdad	1	2	723	996	0	1	930	1600	907.072	204
22	Internal departments for students of Baghdad University/Baghdad	2	1	295	220	0	1	765	650	395.99	123
23	Administrative Building / Ministry of Electricity/Baghdad	1	2	460	385	1	1	765	1300	756.976	190
24	Administrative Building /Council of Governors Baghdad/Baghdad	1	7	8283	2528	4	2	1590	9800	7205.39715	787
25	Medina High School/ Taji Beach/Baghdad	1	2	863	533	0	1	985	1700	864.573	208
26	Administrative building /Council of Governors AL-Anbar/AL-Anbar	3	4	1620	1036	2	2	985	3400	2454.04215	456
27	Administrative building/Council of Governors Nineveh /Nineveh	3	3	1359	766	2	1	985	2550	1698.39	353
28	Administrative building/Council of Governors Salah Eddin/ Salah Eddin	3	5	6321	1805	4	2	1777	7850	5759.73615	731
29	Administrative Building /Ministry of Oil /Baghdad	1	6	7109	2185	4	2	1489	7850	5613.8334	659
30	Directorate of AL-Awqaf Nineveh/Nineveh	3	5	1149	1077	4	2	1777	7850	1774.3005	731
31	Administrative Building/Ministry of Municipalities/Baghdad	1	7	1575	1487	3	2	1284	7850	1981.57365	711
32	Administrative building for the University of Basra/Basra	1	4	987	872	2	2	2209	7850	1217.26185	560
33	An educational building in the Faculty of Medicine of Al-Kindi Baghdad/Baghdad	1	3	1513	610	1	2	892	2295	1470.26355	283
34	Internal departments for students of Diyala University/ Diyala	2	6	2885	1217	2	2	892	4590	3082.2435	584
35	Administrative building of Babylon University / Babylon	1	5	1865	1016	2	2	837	3575	2254.33425	439
36	Services building for the Ministry of Oil/Baghdad	1	2	413	385	1	1	578	960	557.589	175

37	Administrative building of the Ministry of oil/Waest	2	6	1209	1088	2	2	578	2880	1926.9201	511
38	Administrative building for the Ministry of Labor and Social Affairs/Wasit	2	5	1068	915	1	2	534	2200	1495.8027	422
39	Administrative building of the Ministry of oil/ Salah Eddin	3	4	928	717	1	2	567	1880	1392.3546	382
40	Residential units for doctors/Baghdad	1	3	1235	1117	2	2	1051	2730	1697.97495	302
41	Immigration and Displaced Building/ Wasit	2	4	1068	1140	2	2	710	2400	1709.4084	373
42	Internal departments for students of Baghdad University/Baghdad	1	5	1268	1415	2	2	710	3000	1873.13385	416
43	Administrative building of the Ministry of oil /Nineveh	3	2	1038	1196	1	1	1370	2400	1715.766	287
44	Directorate of AL-Awqaf /AL-Anbar	3	5	1828	2116	3	2	1040	4500	3445.0983	571
45	Internal departments for students of University of Karbala/Karbala	2	4	1605	1709	1	2	1040	3600	2441.6721	427
46	Fatima Al - Zahra Secondary School/Waset	2	2	241	243	0	1	322	494	307.585	170
47	Directorate of AL-Awqaf / Baghdad	1	5	1279	1140	1	2	593	2470	1554.1911	395
48	Administrative building of the Ministry of Finance/Baghdad	1	4	1479	1360	2	2	865	2964	1923.39945	362
49	Administrative building of the Ministry of Interior	3	4	2440	2244	3	2	1394	4888	3695.29965	528
50	Internal sections for officers and security associates/Ministry of Interior / Baghdad	1	2	1687	1701	0	1	1952	3458	1956.013	287
51	Abdullah bin Rawahah School/Baghdad	1	2	326	230	0	1	387	612	335.235	160
52	Administrative building of the Ministry of Interior/Baghdad	1	4	1087	689	1	2	589	1960	1214.8521	321
53	Administrative building of the Ministry of Interior/Kirkuk	2	5	1321	850	1	2	589	2450	1582.87395	433
54	Administrative building of the Ministry of Municipalities/Salah Eddin	3	4	1848	1396	2	2	966	3332	2601.17445	452
55	Administrative building of the Ministry of Municipalities/AL-Anbar	3	6	2255	1433	2	2	829	4248	3177.9678	621
56	Administrative building of the Ministry of Municipalities/Najaf	1	6	2217	1305	2	2	754	3840	2322.09285	502
57	Administrative building of the Ministry of Municipalities/Babylon	2	5	2947	1950	2	2	1106	4800	3250.77585	536
58	Administrative building of the Ministry of Municipalitie / Ninava	3	5	5652	3315	4	2	1849	8175	6263.49045	746
59	Administrative building of the Ministry of Municipalities/ Wasit	2	4	720	480	1	2	402	1280	899.02575	323
60	Administrative building of the Ministry of Interior/Baghdad	1	5	1350	910	1	2	578	2400	1421.4354	392

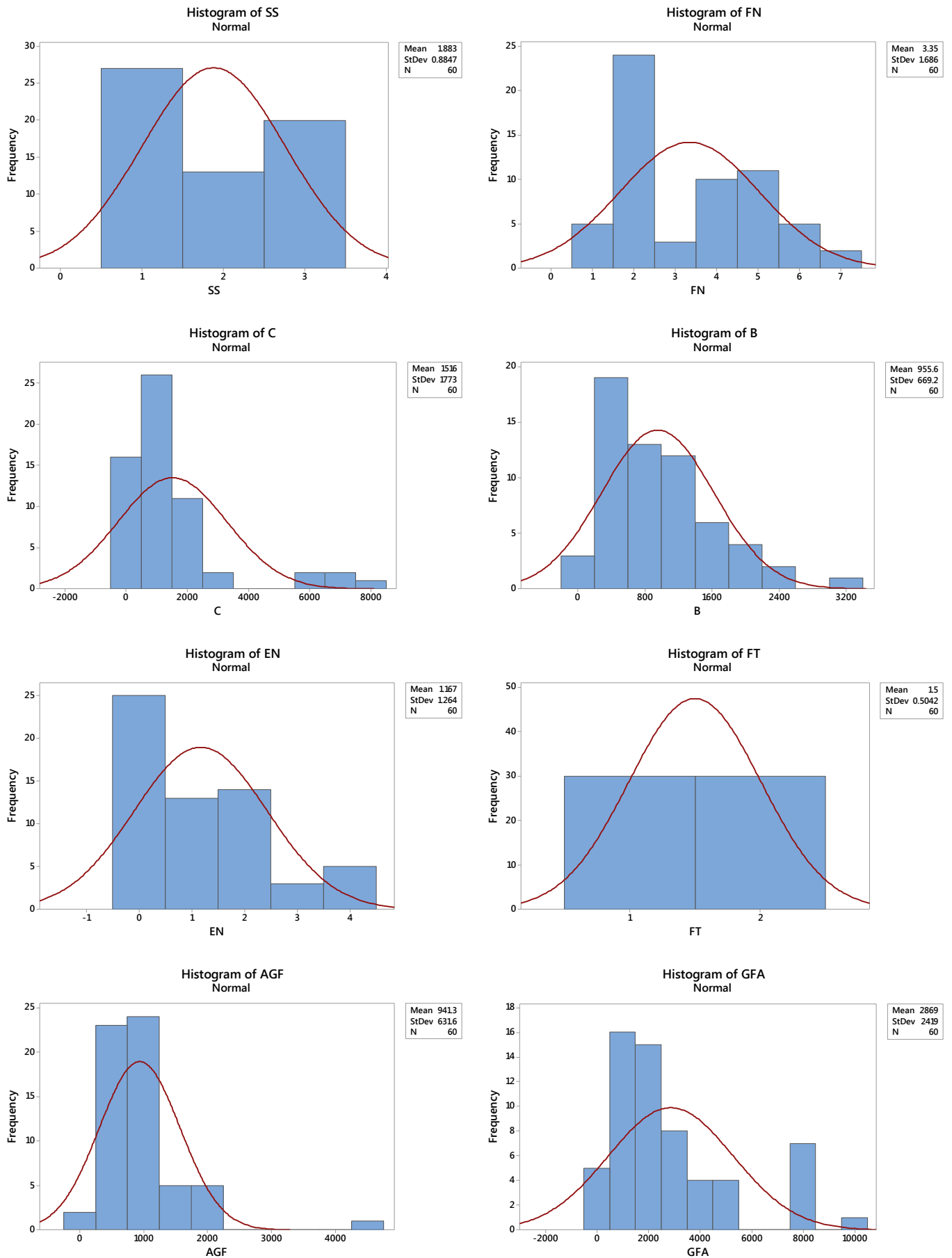


Figure 3. Histograms of independent input variables

Table 5. Descriptive statistics of the variables used in the model development

Description	Mean	Standard Deviation	Range	Min.	Max.
SS	1.88	0.885	2	1	3
FN	3.35	1.686	6	1	7
C	1515.83	1773.455	8262.5	20.5	8283
B	955.62	669.247	3287	28	3315
NOE	1.167	1.2645	4	0	4
FT	1.5	0.504	1	1	2
AGF	941.27	631.60	4334	116	4450
GFA	2869.47	2419.052	9740	60	9800
Output	Total-Cost	Total cost/duration of construction projects			

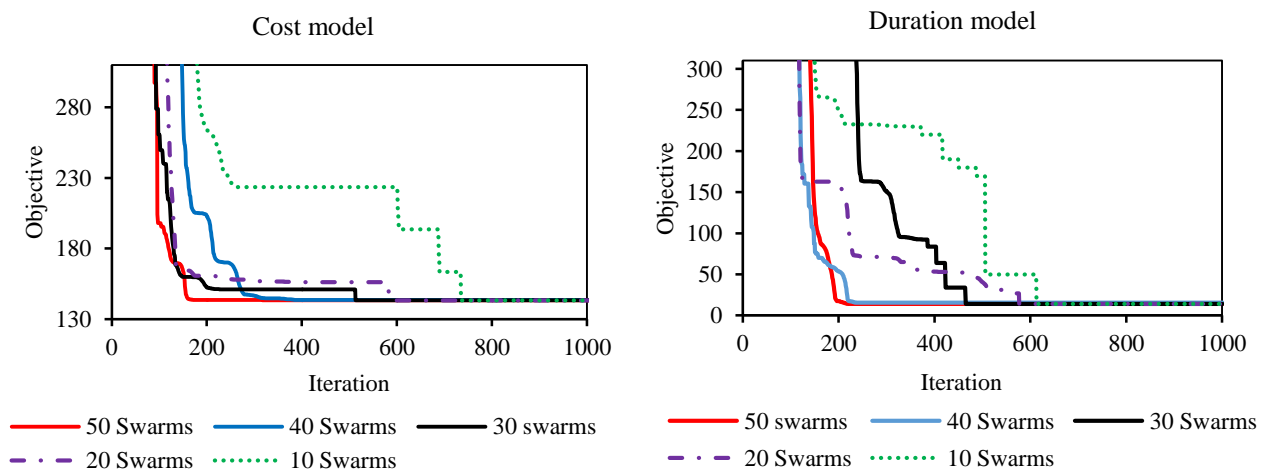
3. Results and Discussion

The PSO technique was utilized to optimize the construction projects cost and-or duration amount. Models have been proposed to examine the influences of swarm size on the outcomes. The main job of the objective function in a PSO approach is to reduce the difference between the predicted and actual cost and/or duration amount. PSO offers models that can assess the cost and-or duration and finding results as close as possible to the measured results. The PSO technique updates its process until either a proper global best (*gbest*) or the maximum epochs (iterations) is achieved, as presented in the methodology. Table 6 shows the parameters used in the PSO model.

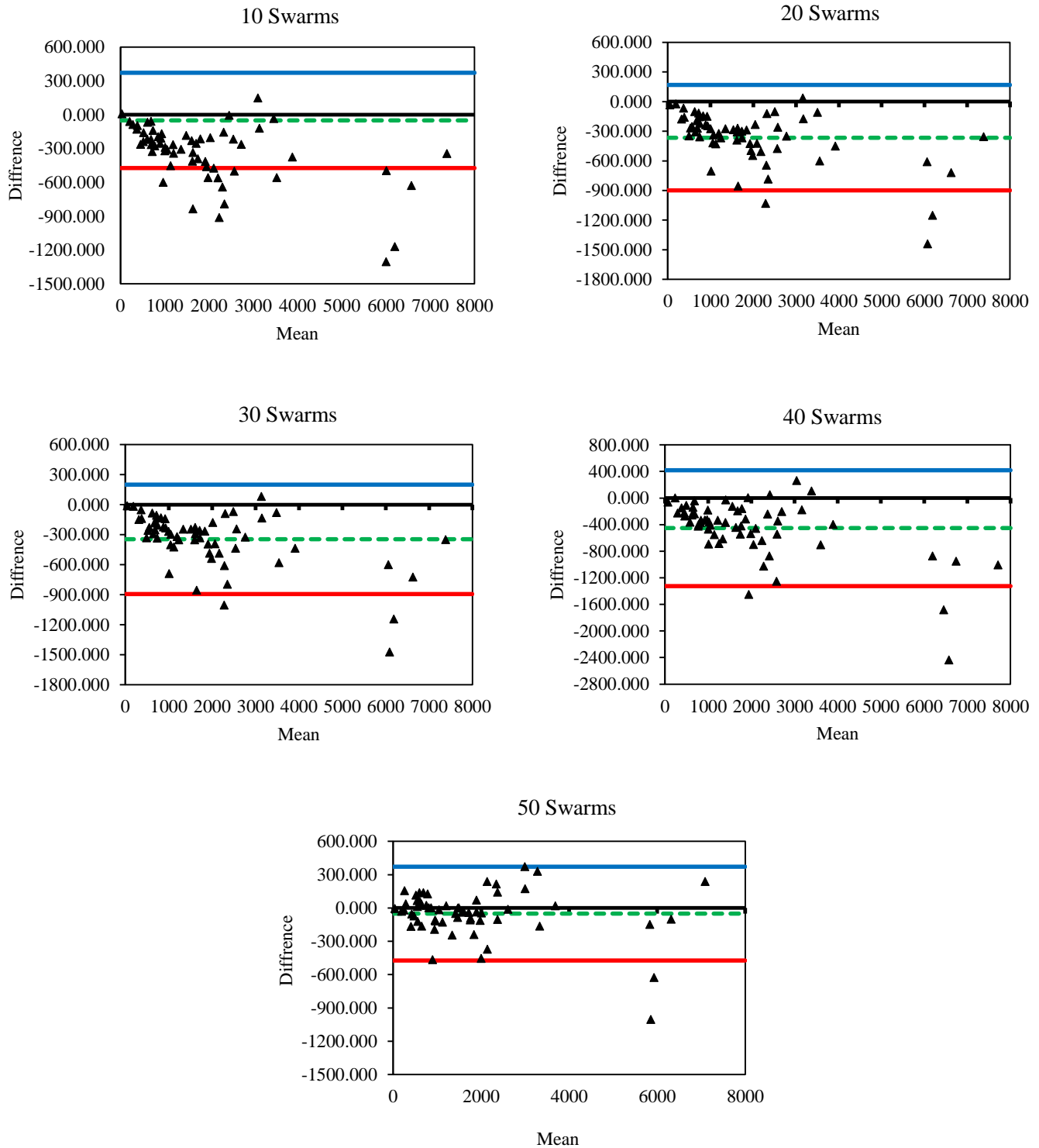
Table 6. Parameters of the PSO

Parameters	Values
Swarm size	10, 20, 30, 40, and 50
Target error	1e-05
Iteration	10000
C_1	1.495
C_2	1.495
w	0.7

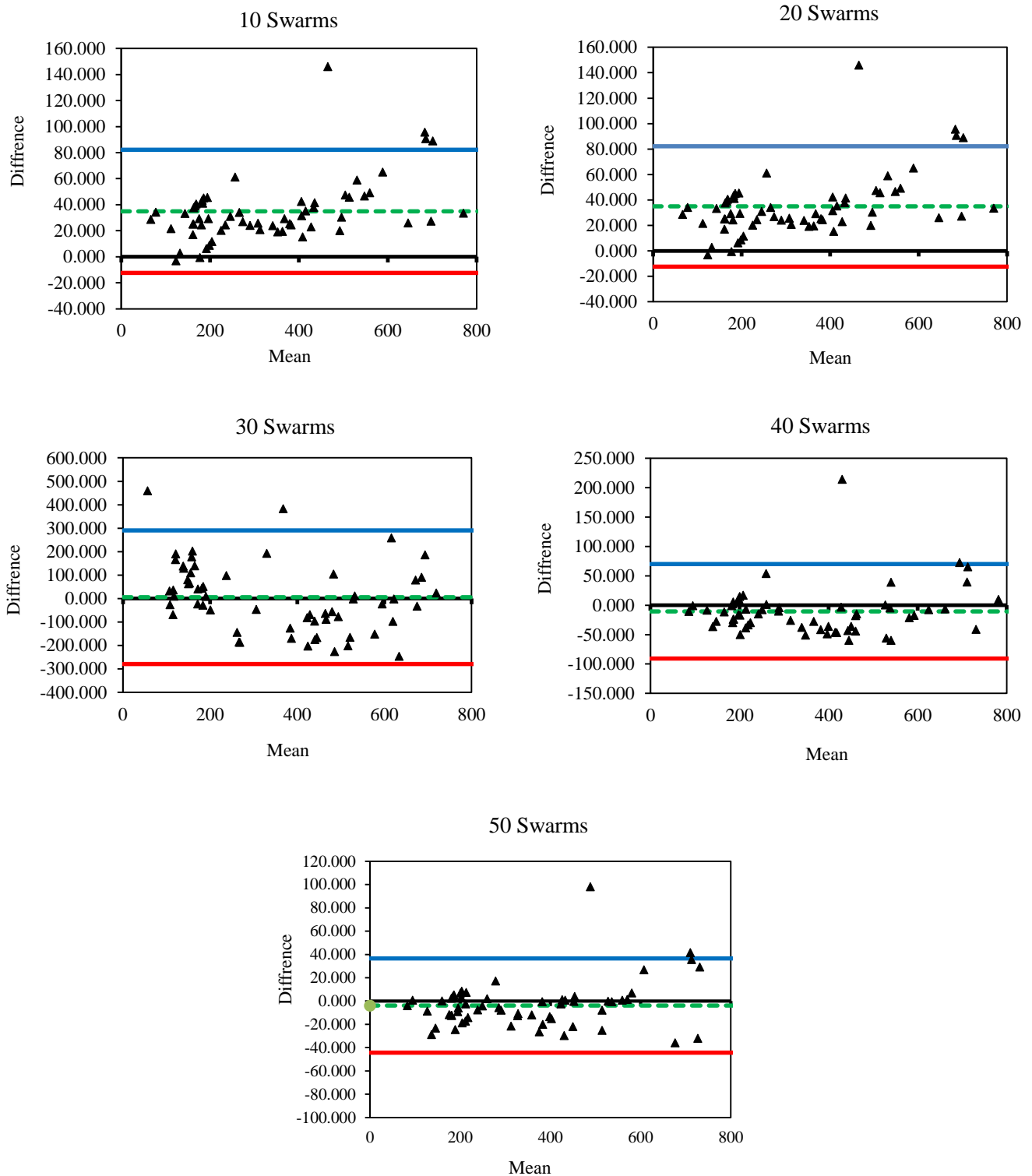
Statistical methods, namely: the coefficient of variation (CoV), correlation coefficient (R), and Bland–Altman (2007) [25] analysis, were used in this study to evaluate and examine the ability of the proposed models. The root mean square error (RMSE) was used as an objective function to choose unknown coefficients. Additionally, five swarm sizes (10, 20, 30, 40, and 50) were used and evaluated. In this study, the iterations number fixed to 1000 because of the differences in the objective functions are stabilized after 700 iterations, as shown in Figure 4 for both cost and duration models. Numerous swarm sizes were tested to evaluate which swarms could minimize the error.

**Figure 4. Objective function (RMSE) versus iteration.**

The Bland–Altman (2007) [25] approach was applied to check the agreement between the actual (A) and predicted (B) cost/duration. The difference between the actual and estimated cost/duration ($A-B$) is plotted against their mean value $((A+B)/2)$, as presented in Figure 5. The Bland–Altman plot exposes the variances in the results and presentations the systematic and random variances. The average value (m) of $+1.96$ and standard deviation (SD) of -1.96 are displayed in the diagram. These values are called the “limits of agreement”. Such limits of agreement specify whether the limits are large or small concerning the overall values, where values outside these ranges show low accuracy.



a) Cost model



b) Duration model

Figure 5. Bland–Altman plot of the relationship between actual and predicted cost and duration values for 50 swarms

As previously determined, Bland-Altman analysis predicts the level of difference. Monitoring scattered values can help to find agreement between actual and predicted values. As shown in Figure 5, a reasonable agreement between the test methods was presented. This figure shows that the data is distributed within the limits of the agreement, indicating the appropriate accuracy of the proposed models. Figure 4 illustrates that 50 swarms provided a better solution for the PSO because, they accomplished the minimum objective functions with 95.83% and 97.92% accuracy, for cost and duration, respectively. With respect to the other swarm sizes, 10 swarms produced significant errors. The results show that the 50 swarms displayed a higher accuracy for the actual values for both models, namely: cost and duration.

According to the CoV and R values, the proposed models accomplish minimum error, as presented in Table 7. The best solution for the PSO algorithm because it accomplishes a minimum coefficient of variation, CoV , and maximum value of correlation coefficient (R), as presented in Table 7. Smith (1986) [26] recommended a rational hypothesis to judge the performance of the model by the following criteria:

- If a model gives $|R| > 0.8$, a strong correlation occurs between the forecast and actual values;
- If a model gives $0.2 < |R| < 0.8$, a good correlation occurs between the forecast and actual values;
- If a model gives $|R| < 0.2$, a weak correlation occurs between the forecast and actual values.

Figure 6 displays that the proposed PSO models had an adequate R -values (0.9441 and 0.9940 for cost and duration) and assessed the target values with adequate accuracy. Moreover, the coefficients (F_1 to F_9) and (K_1 to K_9) obtained from the optimization results will be substituted in Eqs. 5 and 6 of the proposed models, as presented in the following final expressions.

$$Cost = -611.346 + 156.283 \cdot SS + 83.98399 \cdot FN + 0.610803 \cdot C + 0.567313 \cdot B + 92.57055 \cdot EN - 0.16582 \cdot TA + 0.610803 \cdot AG + 0.567313 \cdot FT$$

$$Duration = 2.903068 + 25.6335 \cdot SS + 54.38925 \cdot FN - 0.00249 \cdot C + 0.009338 \cdot B + 13.94593 \cdot EN + 0.036325 \cdot TA + 0.00249 \cdot AG + 0.009338 \cdot FT$$

Table 7. Factors used in the PSO-Cost and PSO-duration models setting

Factor	PSO-Cost	Factor	PSO- Duration
	50 Swarm		50 Swarm
F 1	-611.346	K 1	2.903068
F 2	156.283	K 2	25.6335
F 3	83.98399	K 3	54.38925
F 4	0.610803	K 4	-0.00249
F 5	0.567313	K 5	0.009338
F 6	92.57055	K 6	13.94593
F 7	-0.16582	K 7	0.036325
F 8	0.610803	K 8	-0.00249
F 9	0.567313	K 9	0.009338
M	0.988		0.981
SD	0.174		0.058
COV %	17.6		9.0
R	0.9441		0.9940

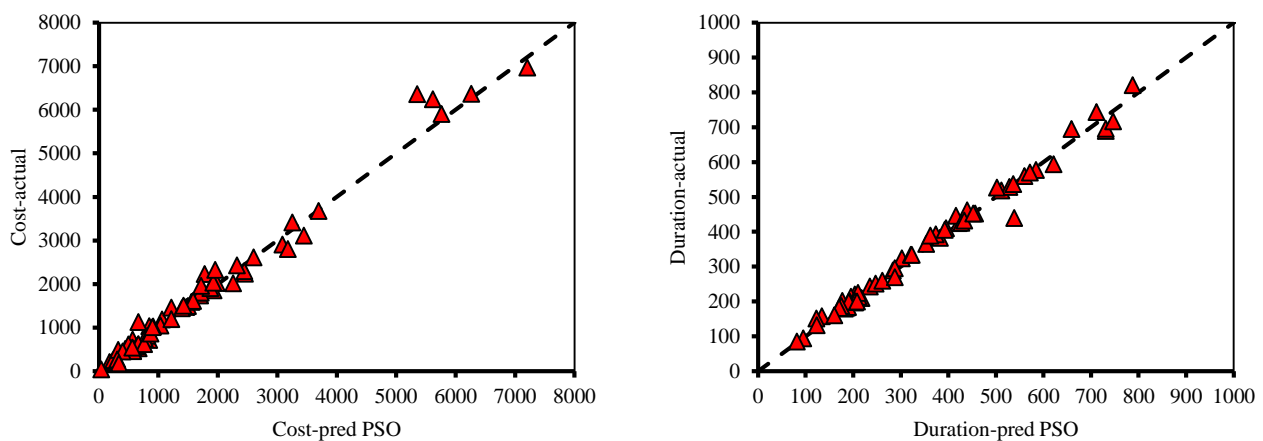


Figure 6. Predicted vs. actual cost and duration values using the proposed model

4. The Validity of the Proposed Models

A dataset comprising of 12 construction projects (20% of the total dataset) was utilized to examine and validate the proposed models. These samples were not utilized in the construction stage of the proposed models. Table 8 shows that the cost and duration assessed by the proposed models are reliable and consistent based on the results. The results recorded values of the mean close to 1.0 (0.97 and 0.99 for cost and duration); this reflected the accuracy of the proposed model, as presented in Table 8.

Pimentel-Gomes (2000) [27] specified that the value of a *CoV* reflects the accuracy of the relationship between the inputs and the output, where *CoV* values of less than 10%, 20–30%, and above 30% mean high accuracy, low accuracy, and low precision, respectively. For the proposed model, the *COVs* for cost and duration models were 10.86% and 4.93%, representing high accuracy. Moreover, the *R*-values of 0.9914 and 0.9940 (as presented in Table 8) reflect a good agreement between the actual and forecast cost and duration values. It can be stated based on these results that the proposed models efficiently assess the cost and duration of the construction projects.

Table 8. Actual database and values predicted using the PSO model

Item	Project ID	Cost actual	Duration actual	PSO-Cost	Cost Actual/predicted	PSO Duration	Duration Actual/predicted
1	Event hall of the Ebad al-Rahman mosque / Yusufiya / Baghdad	177.765	95	207.6	0.86	94.2	1.01
2	School of Habib Ben Khedi/ Nineveh	1058.018	247	1185.2	0.89	251.3	0.98
3	Residential Units for Employees/ Nukhayb /Anbar	237.018	177	254.3	0.93	201.6	0.88
4	Directorate of AL-Awqaf/Kut/ Wasit	768.365	189	743.9	1.03	183.8	1.03
5	Expanding the building of the Faculty of Engineering/University of Baghdad / Baghdad	907.072	204	1022.9	0.89	200.6	1.02
6	Administrative building/Council of Governors AL-Anbar /AL-Anbar	2454.04215	456	2238.2	1.10	452.0	1.01
7	Administrative building of Babylon University/Babylon	2254.33425	439	2015.1	1.12	461.0	0.95
8	Administrative building for the Ministry of Labor and Social Affair /Waset	1495.8027	422	1493.6	1.01	424.6	0.99
9	Administrative building of the Ministry of oil/ Nineveh	1715.766	287	1955.4	0.88	294.9	0.97
10	Directorate of AL-Awqaf / Baghdad	1554.1911	395	1591.8	0.98	410.2	0.96
11	Internal sections for officers and security associates/Ministry of Interior / Baghdad	1956.013	287	2327.8	0.84	269.8	1.06
12	Administrative Building of the Ministry of Municipalities/AL-Anbar	3177.9678	621	2806	1.13	594.1	1.05
				M	0.97	0.99	
				SD	0.105	0.049	
				CoV %	10.86	4.93	
				R	0.9914	0.9940	

The criteria suggested by Golbraikh et al. (2002) [28] were checked for the external verification of the proposed models on the testing datasets. It seems that at least one slope of regression lines (k or k') through the origin should be close to 1.0. Roy and Roy (2008) [29] introduced a confirmative indicator of the external predictability of models (R_m). For $R_m > 0.5$, the condition is satisfied. The squared correlation coefficient (through the origin) between predicted and experimental values (R_o^2) should be close to 1.

The considered validation criteria and the relevant results obtained by the model are presented in Table 9. As can be seen, the proposed models satisfy the required conditions. The external validation criteria result for the models are presented in Table 9.

Table 9. Statistical parameters of the PSO models for external validation

Item	Formula	Condition	Cost model	Duration model
1	$R = \frac{\sum_{i=1}^n (EA_i - \overline{EA_i})(EE_i - \overline{EE_i})}{\sqrt{\sum_{i=1}^n (EA_i - \overline{EA_i})^2 \sum_{i=1}^n (EE_i - \overline{EE_i})^2}}$	$R > 0.8$	0.9914	0.9940
2	$k = \frac{\sum_{i=1}^n (EA_i \times EE_i)}{EA_i^2}$	$0.85 < k < 1.15$	1.0592	1.0340
3	$k1 = \frac{\sum_{i=1}^n (EA_i \times EE_i)}{EE_i^2}$	$0.85 < k1 < 1.15$	0.9441	0.9671
4	$R_m = R^2 \times (1 - \sqrt{ R^2 - R_o^2 })$	$R_m > 0.5$	0.8690	0.7858
where $R_o^2 = 1 - \frac{\sum_{i=1}^n (EE_i - EA_i^o)^2}{(\sum_{i=1}^n (EE_i - \overline{EE_i}))^2}$, $EA_i^o = k \times EE_i$				

The cost and duration values estimation achieved by the proposed models are illustrated in Figures 7 and 8. The models have acceptable estimation accuracy when the ratio of the actual to estimated values is close to one. As can be presented from Figure 7, the ratio distribution of the actual to estimate values for the proposed PSO model in duration have better estimation accuracy than the PSO model in cost.

For further statistical analysis for the mentioned models, a comparison between the actual and assessed cost and duration values has been illustrated in Figure 8. This Figure shows that the proposed PSO model in the duration is closer to the actual duration of projects than the proposed PSO model in cost.

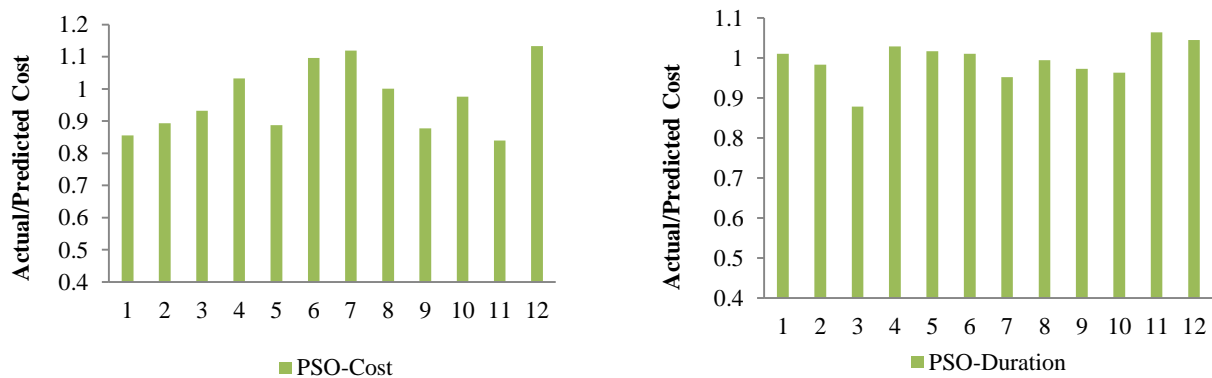


Figure 7. Comparison between the predicted and actual cost and duration amount using the PSO models

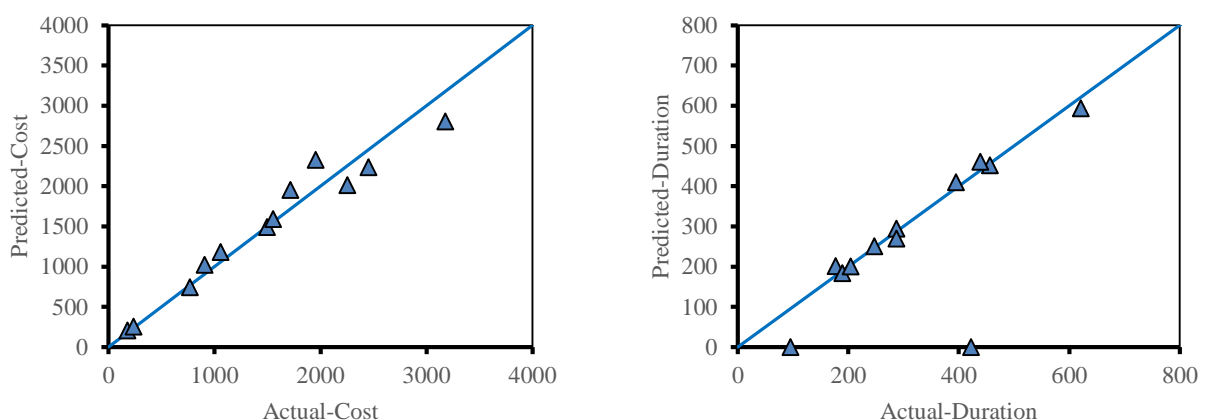


Figure 8. Actual versus predicted cost and duration amount utilizing the PSO models

5. Screening and Parametric Analyses

After constructing the proposed model, various phases were considered: (i) deriving the final models based on collected datasets; (ii) computing several external validation criteria to verify the models; and (iii) conducting a parametric study based on engineering principles and the physics of the problem. The first two steps are purely statistical;

however, the third step is based on engineering principles and should be performed by an engineer who understands the problem being modeled. The first and second steps were achieved here for this type of problem. Therefore, for further verification of the developed model of cost and duration value, a parametric analysis was performed.

This study primarily seeks to assess the effect of individual parameters on cost and duration values. Figure 9 demonstrates the forecast values of the cost and duration accomplished by the proposed models as a function of each parameter. Figure 9 (a) and (b) show the proposed models as a function of the $(C, B, EN, FT, AGF, TFA, FN, SS)$ parameters. Figure 9 (a) and (b) display that increases in the amounts of $C, B, EN, FT, AGF, TFA, FN, SS$ up to a certain level lead to increases in the cost and duration values, indicating that the proposed models can be utilized as a guide to choose the suitable parameters correctly. Moreover, Figure 9 displays that the parameters SS, C , and GFA are the most effect parameters on the cost and duration values.

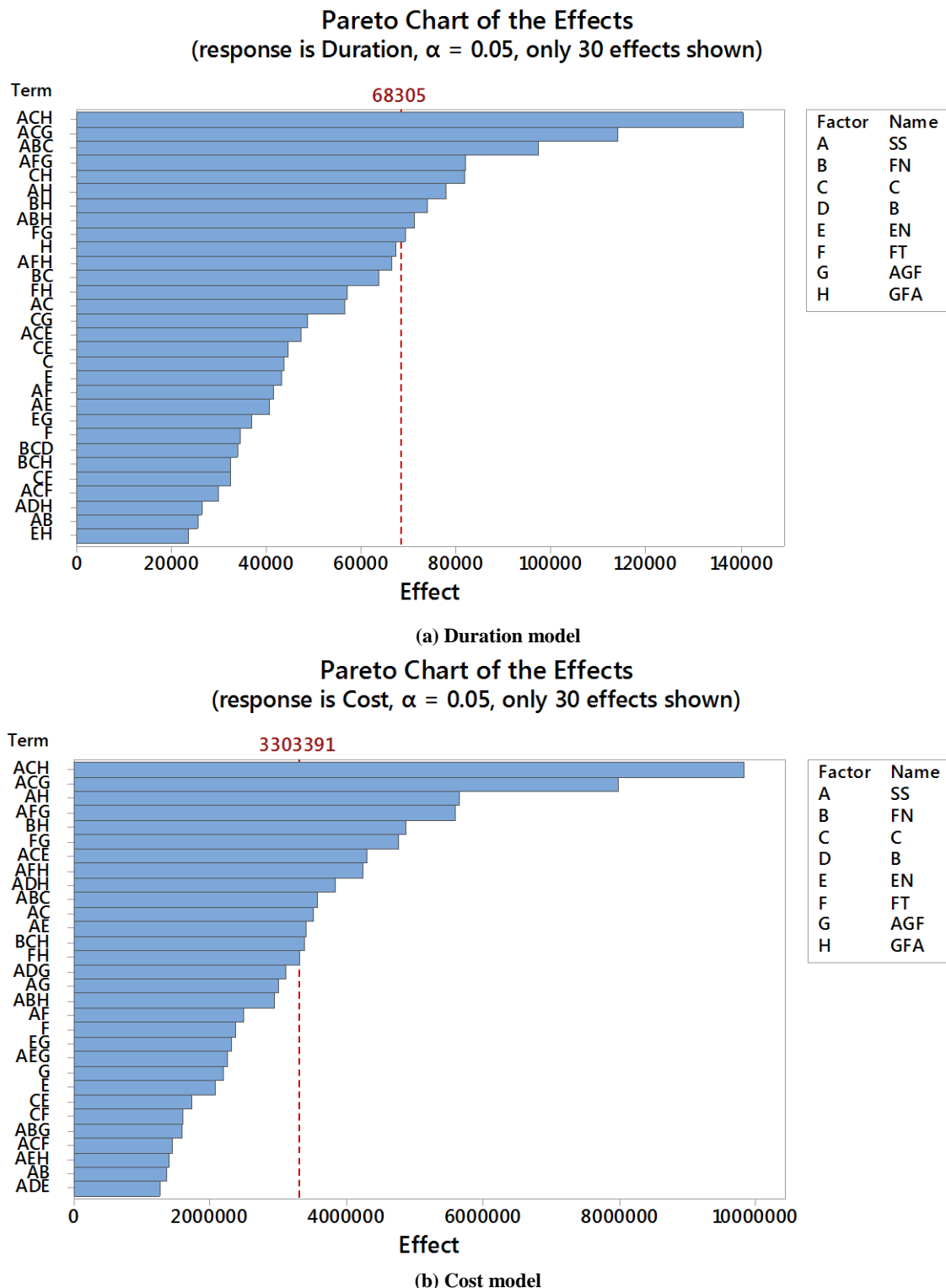


Figure 9 Parametric analysis of the cost and duration values using the proposed models

6. Conclusions and Recommendations

The main objective of this study was to develop mathematical models that will be applied to forecast the cost and duration of the construction projects. In this study, sixty construction projects were utilized to build the proposed models at early-stage design. The main conclusions are drawn according to the models' outcomes, as follows:

- The contractors can utilize the proposed model to assess the construction cost and/or duration, and compare them with that specified by the client at the bid phase, to know if the cost and/or duration will be reasonable for the given project and its budget. This modeling technique based on historical datasets collected from existing projects. Thus, it is more practical, consistent, and reliable than currently utilized subjective methods based on intuitive assessments by designers.
- The statistical analysis demonstrations that the *CoV*, mean, and *R* display good accuracy and reliability for the predicted values. With low mean (0.97 and 0.99) and *CoV* (10.87% and 4.94%) values, the proposed PSO models (for both cost and duration) provide a proper assessment of the construction projects. Hence, this model can be utilized as a design indicator of cost and duration estimations at the early-stage design.
- The outcomes display that the PSO technique is proper for evaluating project management problems and can be utilized as a useful tool to search the optimal solutions with differs parameters.
- The proposed model supplies a guide for choosing the suitable parameters that influence the cost and duration parameters, such as security status, total area, area of the ground floor, floors number, the brick and concrete volume, and elevators number.

In this study, the dataset for only sixty construction projects was utilized to build the model. Nevertheless, more case studies with similar kinds of projects will supply more consistent results.

- Further construction projects should be conducted to examine and modify the proposed model and to investigate a wide range of parameters.
- Future research could be considered to build a model for cost and/or duration estimation for green buildings.

7. Conflicts of Interest

The authors declare no conflict of interest.

8. References

- [1] Foussier, Pierre Marie Maurice. *From Product Description to Cost: A Practical Approach: Volume 1: The Parametric Approach*. Springer Science & Business Media, (2006).
- [2] Layer, Alexander, Erik Ten Brinke, Fred Van Houten, Hubert Kals, and Siegmars Haasis. "Recent and Future Trends in Cost Estimation." *International Journal of Computer Integrated Manufacturing* 15, no. 6 (January 2002): 499–510. doi:10.1080/09511920210143372.
- [3] Rardin, Ronald L., and Ronald L. Rardin. *Optimization in operations research*. Vol. 166. Upper Saddle River, NJ: Prentice Hall, (1998).
- [4] Chinneck, John W. "Practical optimization: a gentle introduction." *Systems and Computer Engineering*, Carleton University, Ottawa, (2006). Available online: <http://www.sce.carleton.ca/faculty/chinneck/po.html>.
- [5] Van Den Bergh, Frans. "An analysis of particle swarm optimizers." PhD diss., University of Pretoria, (2007).
- [6] Bromilow, F. J. "Measurement and scheduling of construction time and cost performance in the building industry." *The Chartered Builder* 10, no. 9 (1974): 57-65.
- [7] Carr, Robert I. "Simulation of construction project duration." *Journal of the Construction Division* 105, no. 2 (1979): 117-128.
- [8] Wang, Yu-Ren, Chung-Ying Yu, and Hsun-Hsi Chan. "Predicting Construction Cost and Schedule Success Using Artificial Neural Networks Ensemble and Support Vector Machines Classification Models." *International Journal of Project Management* 30, no. 4 (May 2012): 470–478. doi:10.1016/j.ijproman.2011.09.002.
- [9] Hong, Yuan, Haibo Liao, and Yazhi Jiang. "Construction Engineering Cost Evaluation Model and Application Based on RS-IPSO-BP Neural Network." *Journal of Computers* 9, no. 4 (April 1, 2014). doi:10.4304/jcp.9.4.1020-1025.
- [10] Zima, Krzysztof. "The Case-Based Reasoning Model of Cost Estimation at the Preliminary Stage of a Construction Project." *Procedia Engineering* 122 (2015): 57–64. doi:10.1016/j.proeng.2015.10.007.

- [11] Lee, Dongoun, Seungho Kim, and Sangyong Kim. "Development of Hybrid Model for Estimating Construction Waste for Multifamily Residential Buildings Using Artificial Neural Networks and Ant Colony Optimization." *Sustainability* 8, no. 9 (September 1, 2016): 870. doi:10.3390/su8090870.
- [12] Juszczuk, Michał, Agnieszka Leśniak, and Krzysztof Zima. "ANN Based Approach for Estimation of Construction Costs of Sports Fields." *Complexity* 2018 (2018): 1–11. doi:10.1155/2018/7952434.
- [13] Hegazy, T., P. Fazio, and O. Moselhi. "Developing practical neural network applications using back - propagation." *Computer - Aided Civil and Infrastructure Engineering* 9, no. 2 (1994): 145-159. doi:10.1111/j.1467-8667.1994.tb00369.x.
- [14] Hanoon, Ammar N., M. S. Jaafar, Farzad Hejazi, and Farah N.A. Abdul Aziz. "Energy Absorption Evaluation of Reinforced Concrete Beams Under Various Loading Rates Based on Particle Swarm Optimization Technique." *Engineering Optimization* 49, no. 9 (December 2016): 1483–1501. doi:10.1080/0305215x.2016.1256729.
- [15] Hanoon, Ammar N., M.S. Jaafar, Farzad Hejazi, and Farah N.A. Abdul Aziz. "Strut-and-Tie Model for Externally Bonded CFRP-Strengthened Reinforced Concrete Deep Beams Based on Particle Swarm Optimization Algorithm: CFRP Debonding and Rupture." *Construction and Building Materials* 147 (August 2017): 428–447. doi:10.1016/j.conbuildmat.2017.04.094.
- [16] Mir, Mahdi, Majid Kamyab, Milad Janghorban Lariche, Amin Bemani, and Alireza Baghban. "Applying ANFIS-PSO Algorithm as a Novel Accurate Approach for Prediction of Gas Density." *Petroleum Science and Technology* 36, no. 12 (March 22, 2018): 820–826. doi:10.1080/10916466.2018.1446176.
- [17] Venkataiah, V., Ramakanta Mohanty, J. S. Pahariya, and M Nagaratna. "Application of Ant Colony Optimization Techniques to Predict Software Cost Estimation." *Computer Communication, Networking and Internet Security* (2017): 315–325. doi:10.1007/978-981-10-3226-4_32.
- [18] Eberhart, Russell C., Yuhui Shi, and James Kennedy. *Swarm intelligence*. Elsevier, (2001).
- [19] Shi, Yuhui, and Russell Eberhart. "A modified particle swarm optimizer." In 1998 IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360), IEEE, (1998): 69-73. doi:10.1109/ICEC.1998.699146.
- [20] Al-Sulttani, Ali O., Amimul Ahsan, Ammar N. Hanoon, A. Rahman, N.N.N. Daud, and S. Idrus. "Hourly Yield Prediction of a Double-Slope Solar Still Hybrid with Rubber Scrapers in Low-Latitude Areas Based on the Particle Swarm Optimization Technique." *Applied Energy* 203 (October 2017): 280–303. doi:10.1016/j.apenergy.2017.06.011.
- [21] Salimi, Shide, Mohammed Mawlana, and Amin Hammad. "Performance Analysis of Simulation-Based Optimization of Construction Projects Using High Performance Computing." *Automation in Construction* 87 (March 2018): 158–172. doi:10.1016/j.autcon.2017.12.003.
- [22] Banyhussan, Qais S., Ammar N. Hanoon, Ali Al-Dahawi, Gürkan Yıldırım, and Ali A. Abdulhameed. "Development of Gravitational Search Algorithm Model for Predicting Packing Density of Cementitious Pastes." *Journal of Building Engineering* 27 (January 2020): 100946. doi:10.1016/j.job.2019.100946.
- [23] Lavanya, Dama, and Siba K. Udgata. "Swarm Intelligence Based Localization in Wireless Sensor Networks." *Multi-Disciplinary Trends in Artificial Intelligence* (2011): 317–328. doi:10.1007/978-3-642-25725-4_28.
- [24] Frank, Ildiko E., and Roberto Todeschini. *The data analysis handbook*. Vol. 14. Elsevier, (1994).
- [25] Bland, J. Martin, and Douglas G. Altman. "Agreement Between Methods of Measurement with Multiple Observations Per Individual." *Journal of Biopharmaceutical Statistics* 17, no. 4 (July 2, 2007): 571–582. doi:10.1080/10543400701329422.
- [26] Smith, Geoffrey Nesbitt. "Probability and statistics in civil engineering." Collins professional and technical books 244 (1986).
- [27] Pimentel-Gomes, F. "Course of experimental statistics." Piracicaba: FEALQ 15 (2000).
- [28] Golbraikh, Alexander, and Alexander Tropsha. "Beware of Q2!" *Journal of Molecular Graphics and Modelling* 20, no. 4 (January 2002): 269–276. doi:10.1016/s1093-3263(01)00123-1.
- [29] Roy, Partha Pratim, and Kunal Roy. "On Some Aspects of Variable Selection for Partial Least Squares Regression Models." *QSAR & Combinatorial Science* 27, no. 3 (March 2008): 302–313. doi:10.1002/qsar.200710043.