

Research Article

Cost Estimation Model Using Modified Parameters for Construction Projects

Sae-Hyun Ji ¹, Joseph Ahn ², Hyun-Soo Lee,³ and Kyeongjin Han⁴

¹Ph.D., Research Associate Professor, Institute of Construction and Environmental Engineering, Seoul National University, Seoul 08826, Republic of Korea

²Assistant Professor, Division of Architecture, Civil and Environmental Engineering, Hoseo University, Asan 31499, Republic of Korea

³Professor, Department of Architecture & Architectural Engineering, Seoul National University, Seoul 151-742, Republic of Korea

⁴Manager, Samsung C&T, Seoul 05288, Republic of Korea

Correspondence should be addressed to Joseph Ahn; josephahn@hoseo.edu

Received 22 February 2019; Revised 26 May 2019; Accepted 3 July 2019; Published 30 July 2019

Academic Editor: Behzad Esmaeili

Copyright © 2019 Sae-Hyun Ji et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Construction projects require huge amounts of capital and have many risk factors due to the unique industry characteristics. For a project to be successful, accurate cost estimation during the design phase is very important. Thus, this research aims to develop a cost estimation model where a modification method integrates influential factors with significant parameters. This study identified a modified parameter-making process, which integrates many influential factors into a small number of significant parameters. The proposed model estimates the cost using quantity-based modified parameters multiplied by their price. A case study was conducted with 24-residence building project, and the estimation accuracy of the suggested method and a CBR model were compared. The proposed model achieved higher overall cost-estimation accuracy and stability. A large number of influence factors can be modified as simple representatives and overcome the limitations of a conventional cost estimation model. The paper originality relates to providing a modified parameter-making process to enhance reliability of a cost estimation. In addition, the suggested cost model can actively respond to the iterative requirements of recalculation of the cost.

1. Introduction

In the production process, only 10 to 15% of the total cost is consumed to design a manufactured good. The rest of 80% of a project cost is designated to where it is needed at this stage. The more a project progresses, the less the possibility there is of controlling the project cost because of the high costs of modifications [1]. Also, construction projects require huge amounts of capital and have many risk factors due to the unique industry characteristics. Therefore, cost estimating plays a significant role during the design stage of a project. For a project to be successful, it is crucial to develop a useful model and system for cost estimation in construction projects [2].

As construction projects keep growing in size and complexity, uncertainties are also increasing, which can be represented by the project cost [3]. Especially in the design

phase, information is prone to change according to the project scope. Clients require rapid and satisfactory responses to all aspects of project management. However, despite the environment of these changing dynamics, previous research focuses on cost estimation models based on similar cases that reflect practical business demands insufficiently.

Moreover, the estimation method tends to become more complicated due to the diversification of cost factors [2], so it is difficult for users who directly perform the cost estimation to apply recent research results to the construction project in practice. This implies that cost models should have the accuracy of prediction, user-centered efficiency, and sustainability that can accommodate cost variance factors.

To deal with these challenging issues, this research aims to develop a cost estimation model where a modification method integrates influential factors with significant

parameters. This study identified a modified parameter-making process, which integrates many influential factors into a small number of significant parameters. Major components for cost model evaluation are presented by analyzing previous research. The development process of a cost estimation model is then presented. A case study was conducted to estimate the cost of reinforced concrete work of residential buildings in Korea. The research outcomes are expected to support decision-making in response to project progress.

2. Literature Review

2.1. Importance of Accurate Cost Estimation. The estimation of project cost plays a key role in the success of a construction project [4]. The process of cost estimating is crucial as it enables construction companies to determine what their direct costs will be and to provide a “bottom line” cost, below which it would not be economical for the work to be carried out [5]. Cost estimation can be described as the process of assessing and predicting the total cost of executing items of work in a given time using all available project information and resources [6]. Estimates are firstly developed at the order-of-magnitude level with an accuracy of -30 to $+50\%$ [7]. They are later refined to the budget and conceptual level with an accuracy of -15 to $+30\%$ and the definitive level with an accuracy of -5 to $+15\%$.

The purpose of cost estimation is to provide appropriate information for decision-making [8]. Overestimated or underestimated cost has the potential to cause lost strategic opportunities to a construction contractor [9]. However, construction contractors continue to use conventional or traditional techniques for cost estimation, such as estimating standard procedures or comparison with similar projects based on documented facts and personal experience. The causes of cost estimation inaccuracy are identified as follows: (1) insufficient time, (2) poor tender documentation, and (3) insufficient tender document analysis [3, 5, 8, 9]. An appraisal framework of cost estimation that can provide a benchmark of activities was suggested.

To overcome the gap between research and practice and evaluate cost estimation models, Carr [8] suggested general estimating principles: realism, level of detail, completeness, documentation, attention between direct and indirect costs and between variable and fixed costs, and contingency. Teicholz [10] asserted the characteristics of desirable forecasting methods as follows: (1) the method should not require input data that are difficult or expensive to collect, (2) the method should be simple enough to permit easy integration into cost systems used for construction projects, and (3) forecasts generated by any desirable method should be accurate, unbiased, timely, and stable. However, there is a trade-off between these criteria: “a forecasting method that is very stable can miss spotting significant trends over the recent past, thus waiting too long before reflecting these trends in the forecast.” This would reduce the accuracy and timeliness of the method. Alternatively, a method that reacts too quickly to recent trends will tend to be unstable and yield sharp changes in forecasts that cannot be trusted. Thus, a balance is needed that can be measured by tests against actual data.

2.2. Cost Model Evaluation Factors. As mentioned, the key issues regarding cost estimation are accuracy, change management, applicability, and data acquisition. Improving the accuracy of cost prediction is generally the purpose of previous studies. Kim et al. [11] compared the performance of cost models between case-based reasoning (CBR) and artificial neural network (ANN) methods. Artificial intelligence (AI) approaches are being employed as an effective strategy [12–21]. The objectives of cost estimation are to predict the accurate bid price and to support successful decision-making. Therefore, accuracy is the fundamental standard of a cost model evaluation [22].

As a project proceeds, the availability of information is increased. Different types of information and various methods are applied at each project stage. As the nature of construction projects changes, new methodologies can be developed. A superior cost estimating model should take various methodologies and different project stages into account. For successful results, a well-rounded model can adapt to internal and external trends. Many resources are required to estimate project cost, including time, information, various personnel, and the cost model itself. These resources are also required to develop, operate, and maintain the cost model. As there are many activities for the resources in the cost estimation process, it is crucial for a cost estimation model to satisfy the different requirements of those who develop and utilize the model. For example, an estimator may not only require an accurate estimate, but also that the model be user friendly. Similarly, operators and model developers might want a model that is efficient and flexible so that it can be easily and effectively managed. Thus, various practical perspectives should be considered when developing a cost model to enhance its applicability.

Since most of the cost models are based on historical data, it is important to acquire sufficient data for estimation. Ellsworth [23] believes that the simplest method to establish a reasonable estimate of facility costs is to identify the costs of similar projects and compare these costs with the cost of the new facility. On the contrary, Kirkham [24] suggests that future construction projects are generally nonrepetitive and tend to have significant and unique features that are likely novel. Such instability can result in estimators not being able to obtain sufficient data for the cost estimation of new, unique projects. As the amount of project data increases, this situation becomes more and frequent.

2.3. Major Components for Evaluation. An ideal cost model cannot be developed using just the actual cost data of a construction but should utilize many other factors, such as individual selection of construction resources, process planning, and the operation sequence. However, for a successful project, it is vital to develop a cost model that can overcome such limitations and to make criteria for a “choice and focus” strategy. Based on the general estimating principles suggested by Carr [8] and the characteristics of desirable forecasting methods identified by Teicholz [10], this research suggests three major components for developing and evaluating cost estimation models: reliability, efficiency,

and flexibility. During the decision-making process, estimators and managers should consider contingency in the project cost due to the difference between research and practice. Therefore, to support confident decision-making, the cost model should be stable and clear. In this study, accuracy, stability, and clarity are considered to be components of reliability.

Estimation is implemented at a given time using all available project information and resources. In many cases, estimators can predict cost with limited resources, while operators and developers should manage cost estimation in an efficient manner without critical endeavor. Efficiency is an important condition for all those who engage in cost estimation because the time for estimating or updating the model is not often sufficient. In this research, efficiency refers to the model simplicity, accessibility, and the inexpensive acquisition of resources.

As construction trends change and technology continues to develop, cost data and new estimation methods will be constantly added and modified. In response to project data additions and deletions, an integrated cost model with flexible characteristics would be very useful because it ensures that separate models are not necessary. Models without flexibility prevent estimators from consistently utilizing the same model. A flexible model can be applied to various project stages, and it is also easy to add data to a flexible model and modify it according to current construction trends.

3. Model Development

This research introduces a cost estimation model based on major components. For the model development, public residential building data from Korea are used. Reinforced concrete (RC) work is selected because it occupies over 40% of the total cost of building construction. The development process is divided into the following six steps: (1) framework selection, (2) estimation method selection, (3) item simplification, (4) influence factor selection, (5) modified parameter making, and (6) database establishment. Figure 1 shows the development process of the cost estimation model. Its framework draws from Porter's value-chain management framework [25].

3.1. Framework Selection. The process framework plays a significant role in maintaining consistency during the model development process. To estimate construction cost, it is crucial to acquire previous cases while constantly obtaining new cases. To ensure that the model is improved and modified continuously, a framework should be chosen that has a cycle, such as the Plan, Do, Check, and Act (PDCA) cycle.

To satisfy the cycle, it is very useful to apply CBR to the cost estimation model. Aamodt and Plaza [26] defines CBR as basically solving a new problem by using information and knowledge from similar previous situation. CBR enhances computer intelligence and makes it more human-like, and there are real values, not artificial [21]. Therefore, it is appropriate to apply CBR to the proposed cost model. The CBR cycle of the highest level may be described by the following

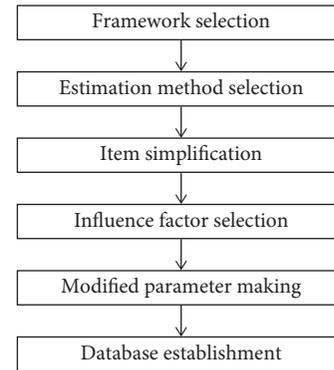


FIGURE 1: Cost model development process.

four processes: (1) retrieving the most similar case or cases, (2) reusing the information and knowledge in those cases to solve the problem, (3) revising the proposed solution, and (4) retaining the parts of this experience that are likely to be useful for future problem solving [26].

From a task-oriented viewpoint, the CBR problem solving mechanism can be described as follows: (1) identifying features, searching, initially matching, and selecting in retrieving, (2) k -NN (Nearest Neighbor) and adapting in reuse, (3) evaluating solutions and repairing faults in revision, and (4) integrating, indexing, and extracting in retention. In this process, retrieve tasks are controlled by influential factors, and reuse tasks are decided by estimation methods. Thus, the database covers all the processes and tasks because it involves all general knowledge involved in all of the previous cases of historical data (Figure 2). From this point of view, this research uses the CBR method as a cost model framework.

3.2. Estimation Method Selection. There are two unit cost application methods: (1) a method based on cost and (2) a method based on quantities. The cost-based method multiplies quantities and unit cost for estimating, and the quantities-based method multiplies unit cost after prediction of the quantities. The method based on cost is suitable for the early design stage and for works in which it is difficult to predict quantities because there are many items. However, this method is not suitable for works which have a changeable unit cost, and escalation needs to be applied.

The method based on quantities is suitable for works which have a small number of items and for the schematic design stage, which has an abundance of information. This method can also be easily used during the definitive design stage and does not require escalation to be applied. Generally, when the cost-based method is used, the efficiency remains higher than the reliability. The application of the quantity-based method yields an opposite effect. Because these methods both have advantages and disadvantages, it is appropriate to apply either one method or a combination of the two methods, depending on the conditions and particular cost-estimating situation.

3.3. Item Simplification. Item simplification is divided into five steps (Table 1), and it starts with analyzing the bill of

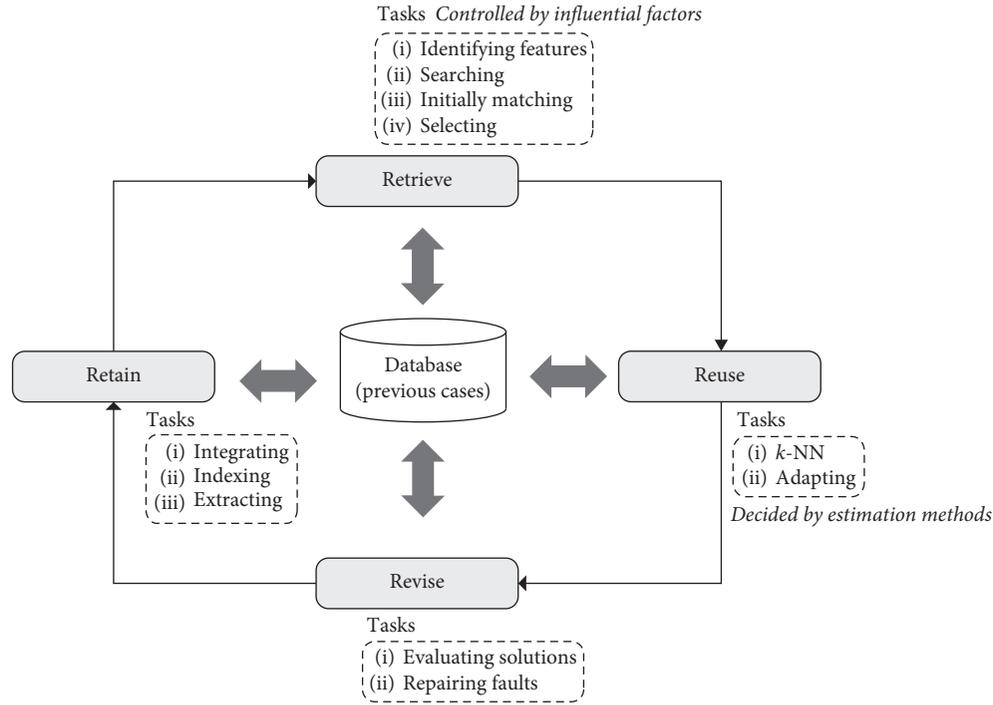


FIGURE 2: Relationship between CBR process and model tasks.

TABLE 1: Item simplification.

Step	Contents	Method	No. of items	Specifications
1	Investigating the no. of items	Analyzing bill of quantities	24	—
2	Grouping items	Concrete, forms, steel bar, and other	24	No. of items: concrete (5), forms (6), steel bars (7), and other (3)
3	Selecting a unit cost application	Concrete, forms, and steel bar: quantity based; other: cost based	21	Cost ratio: concrete 24.88%, forms 49.73%, steel bar 25.17%, and other 0.22%
4	Item simplification	Integrating items which have similar unit costs	15	Ratio of unit cost error: 180 concrete (2): 1.42%, steel bar (6): 1.56%
5	Exception of items estimated by automation	Concrete pouring and steel cutting	12	Concrete pouring = concrete \times 1.01 Steel cutting = steel bar/1.03

quantities (BOQ). The RC work of residential buildings consists of 24 items, and the numbers are usually smaller than those of other works. The items are divided into four groups: concrete (5), forms (6), steel (7), and “other” (3). The total cost of the “other” group has a very small percentage in RC work (0.22% on average). Therefore, instead of estimating “other” directly, it is more efficient to first estimate the other groups and to use equation (1) for the total cost:

$$\text{total cost} = \text{the cost of concrete, forms, and steel} \times \frac{100}{99.78} \quad (1)$$

Items which have similar unit costs are then integrated. In this case, although the accuracy of the reliability decreases a little, the efficiency of estimating is improved. Lastly, items which have no quantities, such as concrete pouring and steel cutting, are calculated automatically by quantities of concrete and steel, so they are separated from the other items.

At the item simplification stage, as items become simpler, the efficiency is improved naturally due to the decrease in the number of items. However, the reliability decreases because errors might occur in item integration and the ratio application stage. Thus, for item simplification, it is important to balance between efficiency and reliability in advance.

3.4. Influential Factor Selection. Influential factors of cost are elements which represent a building’s features and impact its cost. A project is defined by influential factors, and it is possible to describe a project more exactly as the number of influential factors increases [27]. However, if the number of influential factors is large, the number of cases for estimating will also be large. Therefore, it is important to select the appropriate number of factors to improve the efficiency of estimation [28]. Table 2 shows the influential factors of Korean residential buildings.

TABLE 2: Influential factors of Korean residential buildings.

Influential factor	Contents
Basement	Matt (all same)
Gross area (m ²)	1508–10318
Building coverage (m ²)	243–914
Units/building	22–56
Units/standard floor	2–10
Floors of base	1–2
Roof type	Slop, flat
Floors	11–15
Cores	1–4
Pilots type	Partial, whole 1st floor
Pilotis	0–8
Building type	Flat, tower, flat + tower
Floor height (m)	2.8 m (all same)
Unit size (m ²)	49, 59, 84, 114
Combination of unit (m ²)	49, 59, 84, 114, 49 + 59, 49 + 84, 59 + 84, 84 + 114
Type of public area	Hall, corridor
Structure type	Wall type (all same)
Level of finish	All same

3.5. Modified Parameter Making. All new projects cannot agree completely with previous projects. Hence, if such a disagreement occurs during the retrieval process of CBR, we must adapt these values to the current case. In this research, a simple statistical regression method is applied for the adaptation of the model, and the process is divided into three steps: (1) influential factor grouping, (2) parameter modification, and (3) statistical analysis.

3.5.1. Influential Factor Grouping. It is necessary to acquire valid data to use the simple regression method. Although there are many data in the database, all the data are not necessarily valid when a project cost is estimated due to the various characteristics of the data. Therefore, to acquire valid data, data grouping is conducted using influential factors. Data grouping classification is determined by how much information is acquired. Figure 3 explains the different levels of data grouping.

Once project information A is acquired, the 1st estimation can be undertaken, and the valid data are ABD, ABE, ACD, and ACE. Subsequently, once C is added, the 2nd estimation has two valid data: ACD and ACE. Lastly, if E is added, the 3rd estimation has only one valid data: ACE. Additionally, if B or C is not decided and E is acquired, the estimation can use ABE and ACE as valid data. Generally, the 1st estimation has the most data, but it also tends to have lower accuracy than the other estimations. While the 3rd estimation has the least amount of data, it still tends to be more accurate than the other estimations. This indicates that data grouping with a tree structure can be applied to various project stages which have different amounts of project information.

Data grouping helps the cost model's level of flexibility as it helps the model in taking each project stage into account. From the influential factors in Table 2, five major factors are selected for data grouping: unit size, building

type, unit/floor, core, and pilotis. As an example, Table 3 shows the data grouping of 84 m² unit size buildings.

3.5.2. Parameter Modification. After the five major influential factors are selected for data grouping, it is necessary to select parameters (or independent variables) for the simple regression among the influential factors that are left over. As an example, Figure 4 shows the parameter selection of a reinforced concrete item.

Among the 102 historical data, 24 data which have characteristics of 84 m² and 4 units/floor are selected. Figure 4 uses the number of units as a parameter, and statistical effectiveness cannot be guaranteed because the number of units cannot describe all the basement's features. If the modified number of units which cover the basement's features is applied, the statistical effectiveness is improved. The modified number of units is calculated by using the following equation:

$$\begin{aligned} \text{modified number of units} = & \text{number of units} \\ & + \text{number of units per a floor} \\ & \times \text{number of basement floors.} \end{aligned} \quad (2)$$

Table 4 illustrates the relationship between the 12 items of reinforced concrete work and 5 major influential factors, as well as each item's parameters. Column and beam forms use all major factors for data grouping. Stair forms use only two factors (unit size and cores), and the other items except pilotis use four factors. Among the parameters, the modified unit slab includes roof slabs, basement slabs, and unit slabs. The modified unit slab is calculated by using the following equation:

$$\begin{aligned} \text{modified number of slabs} = & \text{number of units} \\ & + \text{number of units per a floor} \\ & \times \text{number of basement floors} \\ & + \text{pilotis of 1st floor.} \end{aligned} \quad (3)$$

3.5.3. Analysis. As more project information is obtained, data groups are subdivided, and the number of cases in the data group decreases more and more (Figure 3). Therefore, to satisfy the reliability of cost estimation using the data group, the statistical significance of the detailed data group should be guaranteed. The statistical effectiveness of the detailed data group should also be higher than that of the superior data group. To analyze and verify these conditions, this research presents an example that has a unit size of 84 m² and 4 units/floor. The results are shown in Figure 5.

Figures 5(b) and 5(c) show the situation to acquire another factor, the building type, under Figure 5(a). Although the number of (b)'s data is smaller than (a)'s, (b)'s coefficient of determination (R squared) is improved from 0.810 to 0.881, and the significance levels of (a) and (b) are almost the same. On the contrary, (c)'s coefficient of

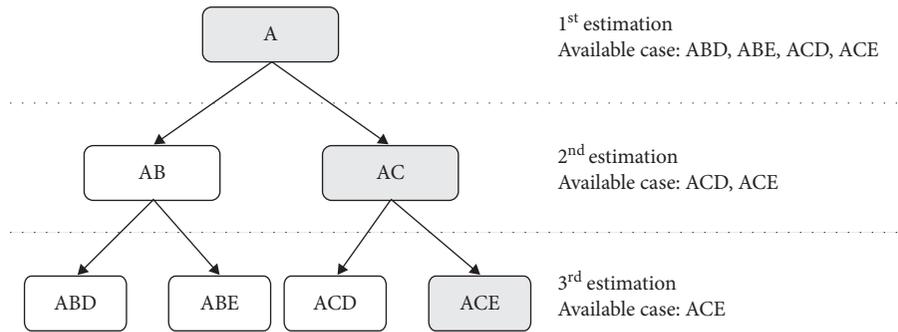


FIGURE 3: Tree structure for concept of case classification.

TABLE 3: Data grouping by influential factors (e.g., 84 m² group).

		84 m ²										
		Flat					Tower					
Unit size (A)												
Building type (B)												
Units/floor (C)		2	4	4	4	5						
Cores (D)		1	2	1	1	1						
Pilotis (E)		2 (2)	1 (2)	2 (2)	2 (4)	1 (4)	2 (2)	2 (4)	1 (4)	2 (2)	2 (4)	1 (5)
Parameter (independent variables)		Grouping					Grouping					

Pilotis (E): the number of pilotis floors (the number of units excepted by pilotis).

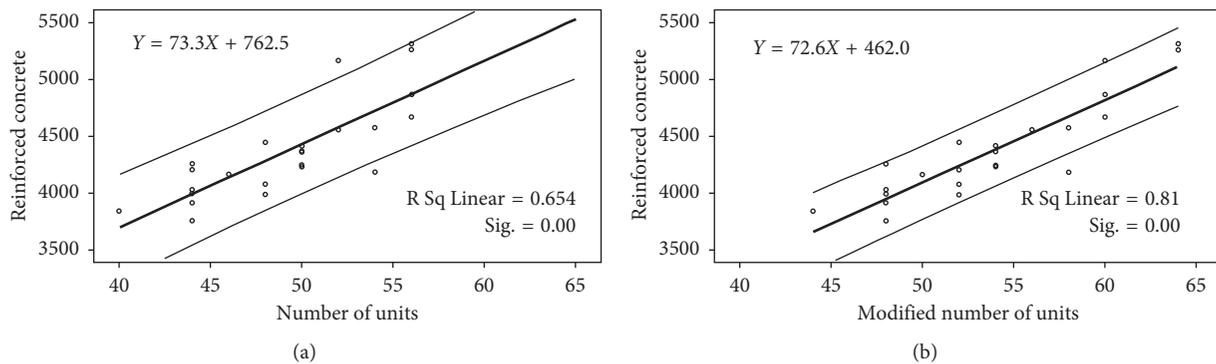


FIGURE 4: Regression analysis (e.g., reinforced concrete).

TABLE 4: Influential factors and parameter of each item.

Items	Influential factors for grouping	Parameter
Reinforced concrete	A, B, C, D	The modified number of units
Concrete	A, B, C, D	—
Steel cutting	A, B, C, D	The modified number of units
Form_basement	A, B, C, D	—
Form_column	A, B, C, D, E	The number of basement floors
Form_beam	A, B, C, D, E	—
Form_flat slab	A, B, C, D	The modified number of slabs
Form_slope slab	A, B, C, D	—
Form_inner wall	A, B, C, D	The modified number of units
Form_outside wall	A, B, C, D	The number of units
Form_stair	A, D	The total number of floors
Form_etc.	A, B, C, D	—

A: unit size; B: building type; C: units/floor; D: cores; E: pilotis.

determination is deteriorated to 0.552, and its significance level is also worse than (a)'s. Through these results, the following was deduced:

- (1) The coefficient of determination and significance levels can help estimators classify a data group logically and statistically, as they allow estimators to

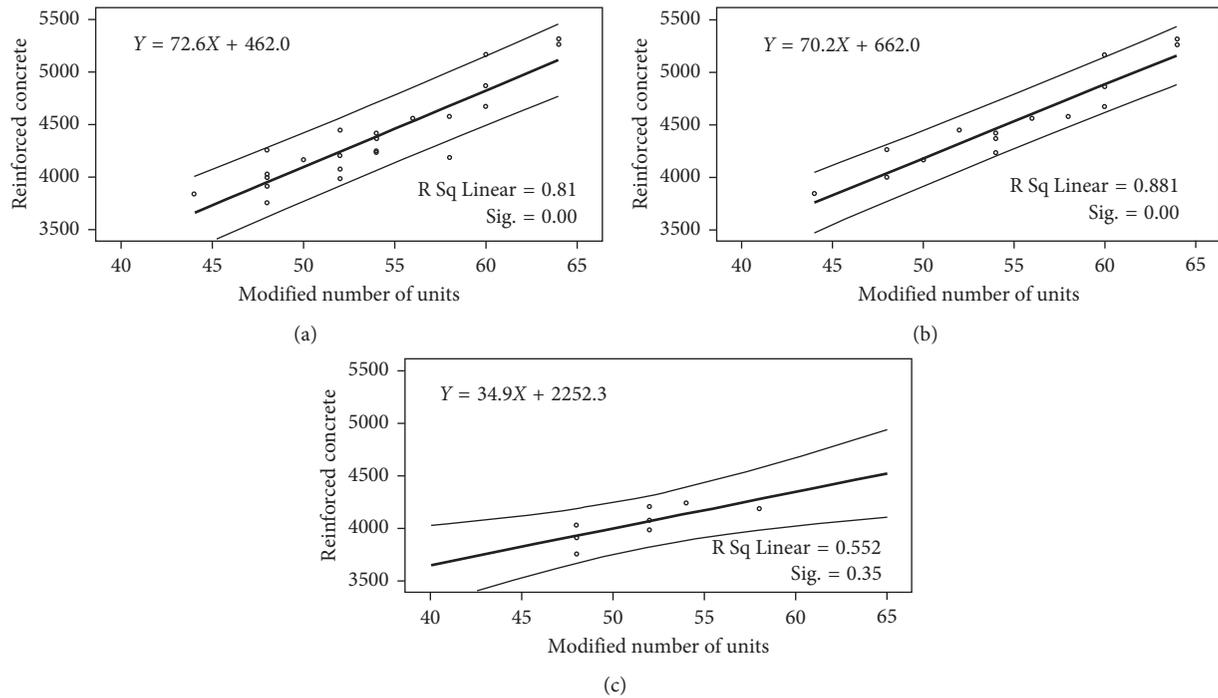


FIGURE 5: Statistical analysis (e.g., reinforced concrete). (a) Data group = 84 m^2 , 4 units/floor. (b) Data group = 84 m^2 , tower, 4 units/floor. (c) Data group = 84 m^2 , flat, 4 units/floor.

consider both the obtained project information and the number of data. For example, after the “tower” building-type factor is obtained in this case, the 84 m^2 -tower-4 units/floor data group (b) will be used for the cost model because it has an improved coefficient of determination and a similar significance level related to the 84 m^2 -4 units/floor data group (a). On the contrary, although the “flat” building type factor is obtained, the 84 m^2 -flat-4 units/floor data group (c), which has 6 cases, will not be used because it yields an opposite result in comparison with (b). Therefore, the 84 m^2 -flat-4 units/floor data group will not be used until the number of data is sufficient for the results to have statistical validity.

- (2) The cost model is qualified as a decision-making tool when the estimating results and basis for contingency are clear. The statistical values of the coefficient of determination and significance level easily guarantee this clarity in the estimating process.

3.6. Database Establishment. The database from this process consists of influential factors, item quantities, item unit costs, and ratios that can be used to estimate other items. In this process, it is most important for the database’s structure to have flexibility, especially if the cost model is to be computerized eventually. Once the database is built with flexibility, it will be possible to add or modify influential factors in response to changing construction trends.

In building a database, it is recommended that a function for additional item simplification be included. This function is required for item simplification (Step 3) and conducting

more approximate cost estimation. For example, 9 form items (Table 4) would be integrated into only 1 item in the early design stage with little information. Third, unit costs should be updated frequently so that estimators can use the latest ones. Finally, ratios for other items estimation should be updated and stored.

4. Validation

Among the major components, the most important factor for decision-making is reliability. Efficiency and flexibility are necessary for mainly using and developing a cost model. Efficiency and flexibility cannot be evaluated until the model is computerized and practically applied, so the verification of the cost model is conducted with a focus on model reliability.

To verify the model, the model results of 3 new cases are compared with the actual results. Table 5 shows the profiles of the validation cases.

k-nearest neighbor (*k*-NN) principle uses a distance measure to classify the *k*-nearest cases in relation to the target case [29]. Then, a selected case is reused to estimate the costs. In this research, 1-NN (the most similar case) is applied and reused to estimate the costs. The 1-NN method is one of the tasks in the CBR reuse process and is compared with the proposed research model. For the cost estimate in the database without a great deal of modification, the 1-NN method just uses the most similar previous case to a new case (Figure 6). Figure 6 is reproduced from Ahn et al. [30]. The results of this comparison are shown in Table 6.

The average error rates in the quantities of this research method (a) and the 1-NN method (b) are 6.28% and 18.72%,

TABLE 5: Profiles of validation cases.

Building name	A		B		C	
	A	a	b	a	b	
Unit size	84	84	114	84	114	
No. of units	58	48	24	42	22	
Units/standard floor	4	4	2	4	2	
No. of elev.	2	1	1	1	1	
No. of floors (ground)	15	13	13	11	12	
No. of floors (underground)	1	1	1	1	1	
Pilotis	×	2 (4)	2 (2)	2 (4)	2 (2)	
Building type	Flat	Tower	Flat	Tower	Flat	

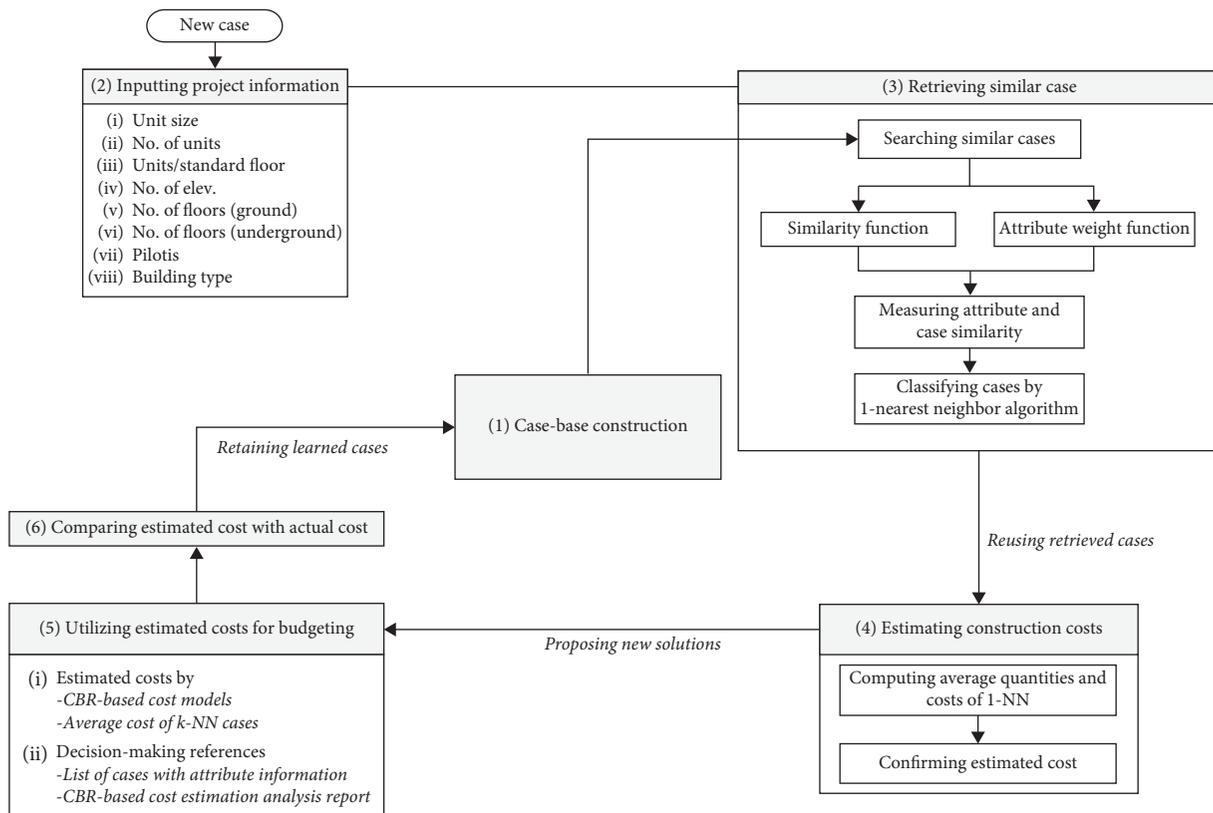


FIGURE 6: The CBR-based cost estimation procedures for validation (revised from Ahn et al. [30]).

respectively, and the average error rates of the total cost are 2.55% and 7.25%, respectively. These results confirm that the proposed method has better accuracy than the 1-NN method. An accuracy rate above 97% satisfies the accuracy requirement.

More cases are required to verify the stability. However, Table 6 shows that method a has better stability in the error rate more than method b. In addition, the stability of items in a case is also as important as the stability of total cases. The total cost in a case is the sum of each item’s cost, which is determined by multiplying the item’s unit cost by its quantity. Even if the accuracy of the total cost is high in a specific case, such a result is hard to be admitted if each item’s error range is too high to be stable. To verify the degree of model stability, the standard deviation of the items’ error rate can be used. A comparison of the standard

deviation of the items’ error rate of the two methods demonstrates that the proposed method has better stability than the 1-NN method.

The proposed model clearly shows how believable the estimation result is because all the items have particular data groups, and the estimator can analyze each group separately. For example, the error rates of “base” and “stair” are lower than all the other items because their statistical values are lower. Through such an analysis, estimators can capture a contingency range of each result.

5. Conclusion

This research developed a cost estimation model where a modification method integrates influential factors with significant parameters. A case study was conducted using

TABLE 6: Validation results.

Building name Method	A		B		C	
	(a)	(b)	(a)	(b)	(a)	(b)
Quantity result						
Reinforced concrete	98.46	96.13	96.56	101.80	97.36	99.46
Concrete	92.49	98.91	100.95	101.08	100.05	93.94
Reinforced concrete pouring	98.46	96.13	96.56	101.80	97.36	99.46
Concrete pouring	92.49	89.91	100.95	101.08	100.50	93.94
Steel bar	86.05	76.23	99.22	113.08	101.50	105.98
Steel cutting	86.05	76.23	99.22	113.08	101.50	105.98
Basement	68.20	52.28	142.55	215.48	147.26	174.87
Column	101.40	107.65	92.09	184.53	124.85	196.76
Beam	94.98	113.91	95.30	153.38	100.97	113.21
Flat slab	96.22	98.41	101.28	105.99	96.49	104.77
Slope slab	95.88	99.91	132.26	183.94	132.26	185.21
Inner wall	97.09	96.86	101.98	102.53	101.60	103.26
Outside wall	95.10	75.99	94.23	95.60	97.41	94.46
Stair	119.16	134.36	116.55	121.50	117.38	123.78
Etc.	0.00	100.00	73.67	0.00	73.67	0.00
Average accuracy (%)	94.43	93.14	104.98	128.21	108.29	121.08
Average error rate (absolute value) (%)	5.57	6.86	4.98	28.21	8.29	21.08
Standard deviation of error rate (%)	10.89	19.62	15.01	39.33	15.87	36.12
Total cost result						
Accuracy (%)	94.20	90.42	98.97	107.71	100.81	104.47
Error rate (absolute value) (%)	5.80	9.58	1.03	7.71	0.81	4.47

(a) Proposed model; (b) 1-NN method in CBR.

residential building data from Korea supported by a government enterprise. As a result, the cost estimation outputs were improved in terms of accuracy and stability. The research findings can be summarized as follows:

- (1) The case study showed that the suggested cost model outperforms a typical CBR cost model
- (2) This study identified the modified parameter-making process, which integrates many influential factors into a small number of significant parameters and has a positive impact on the performance of the cost model
- (3) The proposed model estimates the cost using quantity-based modified parameters multiplied by their price, so the cost model can actively respond to the iterative requirements of recalculation of the cost

The suggested cost model modifies a large number of influential factors into a small number of simple representatives, which can overcome the limitations of conventional cost models. The research outcomes could support and promote cost estimation by decision makers. However, it is necessary to verify the generalization by applying the suggested cost model to various kinds of work and construction projects besides reinforced concrete work for residential projects. The applicability of the modification method and cost modeling process also need to be tested. Future research will require the development and comparison of various cost estimation models and comparative studies with other methodologies besides CBR, such as ANN and conventional parametric methods.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by a grant (19RERP-B082884-06) from Housing Environment Research Program funded by Ministry of Land, Infrastructure and Transport of Korean Government. This research was also supported by the Institute of Construction and Environmental Engineering at Seoul National University and the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (2019R1F1A1058866).

References

- [1] P. Duverlie and J. M. Castelain, "Cost estimation during design step: parametric method versus case based reasoning method," *International Journal of Advanced Manufacturing Technology*, vol. 15, no. 12, pp. 895–906, 1999.
- [2] A. Alqahtani and A. Whyte, "Evaluation of non-cost factors affecting the life cycle cost: an exploratory study," *Journal of Engineering, Design and Technology*, vol. 14, no. 4, pp. 818–834, 2016.
- [3] S. Famiyeh, C. T. Amoatey, E. Adaku, and C. S. Agbenohevi, "Major causes of construction time and cost overruns,"

- Journal of Engineering, Design and Technology*, vol. 15, no. 2, pp. 181–198, 2017.
- [4] A. Jrade and S. Alkass, “Computer-integrated system for estimating the costs of building projects,” *Journal of Architectural Engineering*, vol. 13, no. 4, pp. 205–223, 2007.
- [5] A. J. Smith, *Estimating, Tendering and Bidding for Construction: Theory and Practice*, Springer, Berlin, Germany, 1995.
- [6] A. Kwakye, *Understanding Tendering and Estimating*, Taylor & Francis Ltd., Gower, UK, 1994.
- [7] J. Rast and K. M. Peterson, “Parametric cost estimating for environmental remediation projects,” *AACE International Transactions*, vol. 9, 1999.
- [8] R. I. Carr, “Cost-estimating principles,” *Journal of Construction Engineering and Management*, vol. 115, no. 4, pp. 545–551, 1989.
- [9] A. Akintoye and E. Fitzgerald, “A survey of current cost estimating practices in the UK,” *Construction Management and Economics*, vol. 18, no. 2, pp. 161–172, 2000.
- [10] P. Teicholz, “Forecasting final cost and budget of construction projects,” *Journal of Computing in Civil Engineering*, vol. 7, no. 4, pp. 511–529, 1993.
- [11] S.-Y. Kim, J.-W. Choi, G.-H. Kim, and K.-I. Kang, “Comparing cost prediction methods for apartment housing projects: CBR versus ANN,” *Journal of Asian Architecture and Building Engineering*, vol. 4, no. 1, pp. 113–120, 2005.
- [12] S.-H. An, G.-H. Kim, and K.-I. Kang, “A case-based reasoning cost estimating model using experience by analytic hierarchy process,” *Building and Environment*, vol. 42, no. 7, pp. 2573–2579, 2007.
- [13] J. Cho and J. Chun, “Cost estimating methods for RC structures by quantity takeoff and quantity prediction in the design development stage,” *Journal of Asian Architecture and Building Engineering*, vol. 14, no. 1, pp. 65–72, 2015.
- [14] S. Z. Doğan, D. Arditi, and H. M. Günaydin, “Determining attribute weights in a CBR model for early cost prediction of structural systems,” *Journal of Construction Engineering and Management*, vol. 132, no. 10, pp. 1092–1098, 2006.
- [15] S.-H. Ji, M. Park, and H.-S. Lee, “Cost estimation model for building projects using case-based reasoning,” *Canadian Journal of Civil Engineering*, vol. 38, no. 5, pp. 570–581, 2011.
- [16] S.-H. Ji, M. Park, H.-S. Lee, J. Ahn, N. Kim, and B. Son, “Military facility cost estimation system using case-based reasoning in Korea,” *Journal of Computing in Civil Engineering*, vol. 25, no. 3, pp. 218–231, 2011.
- [17] S. Kim, “Interval estimation of construction cost using case-based reasoning and genetic algorithms,” *Journal of Asian Architecture and Building Engineering*, vol. 11, no. 2, pp. 327–334, 2012.
- [18] C. Koo, T. Hong, C. Hyun, and K. Koo, “A CBR-based hybrid model for predicting a construction duration and cost based on project characteristics in multi-family housing projects,” *Canadian Journal of Civil Engineering*, vol. 37, no. 5, pp. 739–752, 2010.
- [19] C.-W. Koo, T. Hong, C.-T. Hyun, S. H. Park, and J.-O. Seo, “A study on the development of a cost model based on the owner’s decision making at the early stages of a construction project,” *International Journal of Strategic Property Management*, vol. 14, no. 2, pp. 121–137, 2010.
- [20] C. K. Riesbeck and R. C. Schank, *Inside Case-Based Reasoning*, Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 1989.
- [21] N.-J. Yau and J.-B. Yang, “Case-based reasoning in construction management,” *Computer-Aided Civil and Infrastructure Engineering*, vol. 13, no. 2, pp. 143–150, 1998.
- [22] S. Hassim, R. Muniandy, A. H. Alias, and P. Abdullah, “Construction tender price estimation standardization (TPES) in Malaysia,” *Engineering, Construction and Architectural Management*, vol. 25, no. 3, pp. 443–457, 2018.
- [23] R. K. Ellsworth, “Cost-to-capacity analysis for estimating waste-to-energy facility costs,” *Cost Engineering*, vol. 40, no. 6, pp. 27–30, 1998.
- [24] R. Kirkham, *Ferry and Brandon’s Cost Planning of Buildings*, John Wiley & Sons, Hoboken, NJ, USA, 2014.
- [25] M. E. Porter, *Competitive Advantage: Creating and Sustaining Superior Performance*, Free Press, New York, NY, USA, 1985.
- [26] A. Aamodt and E. Plaza, “Case-based reasoning: foundational issues, methodological variations, and system approaches,” *AI Communications*, vol. 7, no. 1, pp. 39–59, 1994.
- [27] M. T. Hatamleh, M. Hiyassat, G. J. Sweis, and R. J. Sweis, “Factors affecting the accuracy of cost estimate: case of Jordan,” *Engineering, Construction and Architectural Management*, vol. 25, no. 1, pp. 113–131, 2018.
- [28] B. Lim, M. P. Nepal, M. Skitmore, and B. Xiong, “Drivers of the accuracy of developers’ early stage cost estimates in residential construction,” *Journal of Financial Management of Property and Construction*, vol. 21, no. 1, pp. 4–20, 2016.
- [29] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, “Miscellaneous clustering methods,” in *Cluster Analysis*, W. A. Shewhart and S. S. Wilks, Eds., John Wiley & Sons, Ltd., Chichester, UK, 5th edition, 2011.
- [30] J. Ahn, M. Park, H. S. Lee et al., “Covariance effect analysis of similarity measurement methods for early construction cost estimation using case-based reasoning,” *Automation in Construction*, vol. 81, pp. 254–266, 2017.

