

## Article

# Preliminary Service Life Estimation Model for MEP Components Using Case-Based Reasoning and Genetic Algorithm

Nahyun Kwon <sup>1</sup>, Kwonsik Song <sup>2</sup>, Moonseo Park <sup>3</sup>, Youjin Jang <sup>4</sup>, Inseok Yoon <sup>3</sup> and Yonghan Ahn <sup>1,\*</sup>

<sup>1</sup> Department of Architectural Engineering, Hanyang University, Ansan 15588, Korea; nhkwon78@naver.com

<sup>2</sup> Department of Civil and Environmental Engineering, University of Michigan, 2350 Hayward St., 2340 G.G. Brown Building, Ann Arbor, MI 48109, USA; kwonsiks@umich.edu

<sup>3</sup> Department of Architecture and Architectural Engineering, Seoul National University, Seoul 08826, Korea; mspark@snu.ac.kr (M.P.); yoon92411@snu.ac.kr (I.Y.)

<sup>4</sup> School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0355, USA; uzjang@gatech.edu

\* Correspondence: yhahn@hanyang.ac.kr; Tel.: +82-31-436-8182

Received: 29 April 2019; Accepted: 29 May 2019; Published: 31 May 2019



**Abstract:** In recent decades, building maintenance has been recognized as an important issue as the number of deteriorating buildings increases around the world. In densely populated cities, building maintenance is essential for ensuring sustainable living and safety for residents. Improper maintenance can not only cause enormous maintenance costs, but also negatively affect residents and their environment. As a first step, the service life of building components needs to be estimated in advance. Mechanical, electrical, and plumbing (MEP) components especially produce many maintenance-related problems compared to other components. In this research, a model was developed that applies the genetic algorithm (GA) and case-based reasoning (CBR) methodologies to estimating the service life of MEP components. The applicability of the model was tested by comparing the outputs of 20 randomly selected test cases with those of retrieved similar cases. The experimental results demonstrated that the overall similarity scores of the retrieved cases were over 90%, and the mean absolute error rate (MAER) of 10-NN was approximately 7.48%. This research contributes to the literature for maintenance management by not only presenting an approach to estimating the service life of building components, but also by helping convert the existing maintenance paradigm from reactive to proactive measures.

**Keywords:** building maintenance; service life estimation; case-based reasoning; genetic algorithm; MEP (mechanical, electrical and plumbing); residential building

## 1. Introduction

In recent years, maintenance has become a key issue of concern for the life cycle of aged buildings around the world [1–3]. Proper maintenance can not only preserve the building performance, but also improve the safety and quality of the residents' lives by allowing deterioration to be recognized in advance [3–5]. In particular, the service life or pattern of building components needs to be carefully examined because it can provide essential information for building management, such as managerial factors and the deteriorated status of components [3,6]. Considering the importance of maintenance for aging residential buildings, being able to estimate when major maintenance will be incurred during a building's life cycle is important [3,7]. In the 1970s, massive numbers of apartment buildings were constructed in South Korea due to urbanization and industrialization [1,3,8].

Such apartment buildings have increased because of the housing supply policy promoted by the government [3] and make up the majority of housing in Korea [9]. According to the Ministry of Land, Infrastructure, and Transport [9], outdated buildings over 30 years old account for 44.6% (3,205,774 buildings) of all buildings (7,191,912 buildings) in South Korea, and the proportion is expected to continue to increase [10,11]. As indicated in Table 1, residential buildings account for 64.3% (4,625,077 buildings) of all aged buildings [9], which may lead to safety-related problems and degrade the original performance. These buildings can experience diverse problems such as concrete neutralization, corrosion of reinforcing steel, and equipment disorder that affect maintenance and commonly begin 10 years after the building is completed [3,10].

**Table 1.** Status of deteriorating buildings depending on the region in Korea.

Classification	Residential	Commercial	Industrial	Factorial	Cultural and Educational	Sum
~10 years	630,503	283,755	98,130	39,635	195,405	1,247,428
~10–15 years	237,776	145,490	51,846	29,535	108,892	573,539
~15–20 years	325,726	157,442	50,188	27,154	89,013	649,523
~20–25 years	412,553	152,590	36,986	29,061	139,960	771,150
~25–30 years	473,771	145,370	30,403	17,034	77,920	744,498
~30–35 years	355,470	94,446	14,011	10,434	22,924	497,285
~35 years	2,189,278	291,871	35,601	41,185	150,554	2,708,489
Sum	4,625,077	1,270,964	317,165	194,038	784,668	7,191,912
(%)	(64.3%)	(17.7%)	(4.4%)	(2.7%)	(10.9%)	(100%)

Especially, mechanical, electrical, and plumbing (MEP) components incur many of the maintenance-related problems of a building. MEP refers to nonstructural building facilities such as electricity, gas, water supply, heating, air conditioning, ventilation, and elevators [3,12,13]. MEP-related maintenance costs may account for 50% of the total life cycle cost (LCC) for a large and complex building [12,14]. MEP is a significant part of buildings because it provides essential functions for the sustainable and comfortable living of residents. Therefore, appropriate maintenance management is required to preserve the original function and performance during its operation. Inadequate maintenance management can not only cause economic damage to residents because of building deterioration but also degrade the quality of the residents' lives [1,15]. As a first step for building management, efforts need to be made to systematically estimate the service life in advance before significant deterioration begins.

Various studies and efforts have focused on estimating the optimal service life of components to determine when the maintenance of a building should be carried out. However, previous research presents limited performance towards estimating the service life from a long-term perspective because of the lack of collected past data and estimation methodologies. Furthermore, there has been limited consideration of various factors such as the building characteristics, maintenance types, number of maintenances, completion year, environment, and maintenance cost. Meanwhile, traditional building maintenance is based on fixed repair cycles of components, where each component is commonly repaired according to a predetermined mean time. However, such a deterministic approach has a limited ability to consider maintenance-related factors or uncertainties of building components. Furthermore, the scope of the maintenance is very extensive to be covered because a building comprises a number of components. This makes it more challenging to estimate the service life for maintenance and consequently deal with potential maintenance-related problems. In this respect, maintenance management based on previous data can be utilized as an alternative or supplement to current estimation approaches. This is because data from past projects have various types of information, such as the gross area, number of floors, number of households, parking lots, area for management, year of completion, and type and number of repairs, which enable reliable estimation. Thus, the present research focused on applying previous data regarding building maintenance to estimation. However, relevant past data are not systematically organized, and the factors associated with the service life of components and building maintenance have been inadequately examined.

To deal with this limitation, a model is developed for estimating the service life for maintenance based on previous data during the preconstruction phase. The scope of this research is limited to examining the service life of MEP components for which building repair occurs on a large scale, which would have a significant effect on the building performance. First, preliminary research is conducted on building maintenance and service life estimation based on an extensive literature review. Then, the limitations of existing approaches are identified. A model for estimating the service life for building maintenance is developed for application to current maintenance management. A database is established based on previous data or cases associated with the service life. Next, attributes are extracted and then weighted with the genetic algorithm (GA). Based on the weighted values of attributes, the similarity among cases is computed from the weighted Euclidean distance (WED). The most similar cases are extracted from the database in order to estimate the service life or pattern. Finally, the estimated values are validated by comparison with the original service life of the test cases. This estimation approach can be used to investigate the maintenance or repair that may be incurred by the aging or deterioration of building components; thus, it is expected to help contractors establish a preventive strategy for building maintenance in advance.

## 2. Preliminary Research

### 2.1. Literature Review

Building maintenance management is essential to ensuring the performance of a building during its life cycle and has been recognized as a major issue as the number of deteriorating buildings gradually increases [3,16]. Thus, proper management is needed to maintain the building performance and ensure residents' safety and needs. However, this is a challenge because a building comprises various components that are affected by different factors such as the lifestyle, surrounding environment, usage pattern, and weather [1,12]. In particular, MEP components are an important part of buildings [12]. MEP components typically comprise diverse subsystems such as heating, ventilation, air conditioning, power distribution, fire protection, water supply, and drainage [3,12,13]. In other words, they are combinations of subsystems complexly connected with equipment, wire, and pipes. Thus, inappropriate maintenance management may cause significant costs or degradation of the building performance. To deal with building maintenance problems, the service life or pattern should be identified as a first step towards establishing preventive measures. Many researchers have devoted a lot of efforts to improve current maintenance management and establishing management plans, as listed in Table 2.

**Table 2.** Approaches and factors highlighted in previous research.

Authors	Research Objective	Target and Scope	Research Methodology	Highlighted Factors	Outcomes
Arif et al. [20]	To propose decision support framework related to infrastructure maintenance	Infrastructure	Mathematical, Probabilistic	Route, age of infrastructure, traffic, capacity, aging condition, safety, accessibility, affordability	Decision support framework
Cho and Yoon [21]	To develop a model for determining cost-effective renovation time	Building renovation	Mathematical	Temperature, energy cost, area, operation period, renovation time	Decision support model
Choi et al. [19]	To determine a maintenance cost based on regression analysis	Highway	Statistical, Regression, and Cluster Analysis	Temperature, pavement age, maintenance cost, precipitation	Maintenance cost
Elcheikh and Michael [22]	To predicting maintenance cost for asset management	Canal system	Probabilistic, Monte Carlo simulation	Asset type, maintenance cost	Maintenance cost
Kim et al. [3]	To evaluate maintenance cost in apartment building	Residential Building	Probabilistic, Monte Carlo simulation	Period after completion, work trade, maintenance cost, area, year of repair, number of households	Maintenance cost distribution
Lee and Ahn [12]	To analyze service pattern in building MEP components	Building's MEP component	Probabilistic, Monte Carlo simulation	Work trade, repair period	Service life distribution
Sharma et al. [17]	To estimate preliminary cost for operation and maintenance	Water treatment plant	Mathematical	Work trade, capacity, construction cost of subsystems	Annual operation and maintenance cost
Stenbeck [18]	To examine the effect of snow on maintenance cost in highway	Highway	Statistical, Mathematical	Weather (climate), maintenance cost, location	Relationship on snowfall and maintenance cost
Park et al. [1]	To examine maintenance period in finishing work	Finishing component	Probabilistic, Monte Carlo simulation	Work trade, repair period	Service life distribution

Park et al. [1], Kim et al. [3], and Lee and Ahn [12] suggested probabilistic methods for estimating the maintenance cost and service life pattern of apartment buildings. Kim et al. [3] used a loss distribution approach (LDA) to create a detailed matrix associated with the risk for maintenance. They thought that it is essential to consider the uncertainty when estimating the maintenance cost. Thus, they applied a Monte Carlo simulation (MCS) to deal with the uncertainty depending on the building conditions and characteristics. Based on the matrix, a distribution was derived for total the maintenance cost. The distribution demonstrated that maintenance 11–20 years after construction is closely associated with the repair cost. Park et al. [1] and Lee and Ahn [12] attempted to analyze service life patterns from a long-term perspective. They also adopted a probabilistic approach based on historical data to identify repair patterns in residential buildings. They focused on examining the service life for MEP and finishing components because these account for the majority of maintenance-related problems. They showed the service lives of building components based on a probabilistic distribution, and confirmed that each component has a different repair pattern depending on its characteristics. Their works are notable in that they suggested predictive methods for identifying maintenance-related costs and service life based on historical data. However, they merely focused on a stochastic approach towards maintenance management, so their ability to validate their results was limited. Furthermore, their research may have limited applicability to building maintenance because they only presented the probabilistic distribution for the maintenance cost or repair period; this is inadequate for site managers to recognize a specific degree of maintenance cost and service life.

Some groups have estimated the maintenance cost by focusing on the infrastructure. Sharma et al. [17] estimated the maintenance cost of water treatment plants during the planning phase. They developed an equation for estimating the maintenance cost based on historical data updated with the construction cost index (CCI). They performed multiple regression analysis on the updated cost to develop a preliminary cost estimation model that includes operation and maintenance costs. They compared the estimated results with the original bidding cost to validate the proposed equation. Their model is a simple and convenient method for estimating the cost, but they did not ensure the accuracy because the cost estimation can change significantly depending on the site conditions, weather, competition among bidders, and economic conditions. Furthermore, the estimation scope addressed by the model is very extensive. These limitations may cause inaccurate results. Stenbeck [18] elaborated upon the correlation between highway and maintenance costs during the winter season. In order to identify the linear relationship, the maintenance costs and snowfall based on weather data were utilized. The results confirmed that the cost and highway stations are correlated. This research is applicable to selecting the optimal location with regard to weather. However, there was limited consideration of various factors related to maintenance because the analysis focused only on snowfall data. Choi et al. [19] developed a model for estimating future maintenance costs based on five factors: the pavement age, temperature, precipitation, thickness, and traffic loading. Statistical approaches, such as cluster analysis and regression analysis, were utilized to derive influential factors with regard to the maintenance cost. They found that the traffic load is the most critical factor that affects pavement degradation. However, some results regarding precipitation are limited because they may be changed depending on the data used in the experiment.

Other groups focused on strategic maintenance management and planning to support decision-making. Arif et al. [20] proposed a framework for supporting decision-making regarding infrastructure maintenance investment. They adopted the multiattribute utility theory (MAUT), Markov decision process (MDP), and portfolio management to derive decision support results. They assessed their framework through case studies of four bridges in terms of three aspects: socioeconomic, infrastructure utilization, and physical condition. The developed framework provides useful data for infrastructure maintenance (i.e., performance trend curves, decision logic maps, and a network-level maintenance investment plan). Cho and Yoon [21] developed a decision support model for estimating the best renovation time in terms of economics. They carried out a case study to identify the optimal cost-effective time. They analyzed five parameters to verify the developed model: the

installation costs, energy costs, benefit, alternative value, and renovation time. The results demonstrated that the renovation time is closely associated with the initial installation cost. Elcheikh and Burrow [22] developed an integrated maintenance model that considers data uncertainty. They performed a probabilistic approach based on MCS to identify the maintenance cost and deal with uncertainty related to building deterioration and repair cost. They found a large variation between preventive maintenance and repair costs, which represents larger uncertainty in assets. Their findings can be utilized in asset management to improve upon existing decision-making associated with cost budgeting and investment. These studies are noteworthy because factors from diverse perspectives were applied to the estimation. However, such approaches are limited at estimating the maintenance-related cost during preconstruction because they address extensive scopes such as economics and assets. Thus, the cases or data for validation are inadequate. This makes reliable estimation for repair cost and validation of the proposed approaches a challenge.

The previous literature can be summarized as follows. Some groups estimated the maintenance cost regarding infrastructure to establish strategies and plans for maintenance management, which can further support decision-making by maintenance managers. However, they covered extensive ranges to determine the maintenance cost or renovation time. Thus, there are insufficient data for estimating the cost or time of specific maintenance. In addition, a limited number of factors were considered. Other groups focused on estimating the maintenance cost and service life of apartment buildings. They used MCS as a primary methodology to deal with the uncertainty of building components. However, their works depended on the probabilistic approach, which has a limited ability to provide the specific maintenance cost or service life to building managers. Furthermore, it is challenging for managers to determine a specific maintenance cost or time for repair because the estimated results represent approximate tendencies according to a probabilistic distribution.

## 2.2. Case-Based Reasoning

In this research, case-based reasoning (CBR) was utilized to estimate the service life of a building component during the preconstruction phase. CBR is an artificial intelligence method that solves a given problem based on data and knowledge retrieved from past similar cases. It stems from cognitive science and is based on how humans think [23–27]. This methodology has been widely adopted in diverse fields such as cost estimation, market selection, safety diagnosis, and building maintenance [28–31] because it can find solutions even when the current problems are well-structured or related data are insufficient or limited [32–34]. Nevertheless, applying CBR to maintenance cost estimation is a challenge because relevant data have not been collected systematically. Therefore, a database needs to be established first. In general, CBR is composed of four phases: case retrieval, reuse, revision, and retention [23,30,34]. Depending on the similarity score, the most similar cases to a given case are extracted. The extracted cases are reused to solve the current problem. If the estimated results are limited or inadequate as a solution, they should be revised to match the problem. Then, the revised results are retained as a new case in the database [35,36]. This research focuses on case retrieval and reuse because they are recognized as the critical phases of CBR [31,35,37,38]. This approach has the following benefits. First, CBR can address various attributes for the estimation, including numerical and nominal data (e.g., areas, building type, number of households, year of completion, and repair items) [28,31]. The results estimated from similar cases are reliable and accurate because they are based on actual cases in the past. CBR enables contractors or site managers to estimate the service life from a long-term perspective. In CBR, cases are retrieved based on similarity, which is determined by the distance measurement function and weights of the attributes [23,34]. Thus, the optimal weights of the attributes need to be determined [39,40]. Because of its importance to the retrieval phase, many researchers have focused on finding the optimal weight for extracting the most similar cases [31,40–43]. Various weighting methods have been considered, such as feature counting (FC), the gradient descent method (GDM), the analytic hierarchy process (AHP), and multiple regression analysis (MRA) [39,40,44]. However, such methods fail to search for the optimal weights or



perform an unnecessary search process, which reduces the reliability of the retrieved cases [42,45,46]. In this research, a genetic algorithm (GA) is utilized because it is recognized as the most effective weighting method [39,41,45]. The attribute weighting is described in Section 3.3.

### 2.3. Genetic Algorithm

In this research, GA is adopted to calculate the weights. This is a heuristic optimization algorithm based on the evolutionary concepts of natural selection and genetics [38,45,47–49]. Principles of biological evolution such as selection, crossover, mutation, and recombination are used to determine the optimal weights for explaining the relationships between the attributes and output [49,50]. GA has been widely applied to optimization problems in business, engineering, and the natural and social sciences because it has a powerful search ability and supports simple computations without derivatives of the objective function, even if a given problem is nonlinear or difficult to solve [38,49,51,52]. In GA, the fitness of a set of genes represents how well a solution satisfies a given problem [38,48]. GA performs a parallel search, while traditional algorithms start a search at one point. During evolution, genes of each generation exchange information accumulated by previous generations and start searching for a new area [50,52]. The search direction is determined probabilistically for each generation. Thus, GA can search for a global optimal solution and reduce the probability of falling into a local optimal solution [49,50]. A GA model sets the candidate solution and fitness function (i.e., objective function) to evaluate the solution. The candidate solution is a set of parameters that represent a proposed solution to the problem. In order to search for the optimal value, the parent population is first selected at random to include a chromosome with a dominant trait. The population size refers to how many chromosomes are in one generation. If there are too few chromosomes, GA has few possibilities for crossover, and only a small part of the search space is explored [45]. If there are too many chromosomes, GA slows down [45,49,50]. The fitness function evaluates strings (chromosomes) in order to select dominant ones that continue into the next generation [38,51]. As illustrated in Figure 1, GA only selects genes suitable to a given problem. The genes are selected and crossed with each other to transfer better genes to the next generation [45,52]. Then, the chromosomes including genes are mutated because the selection of the best genes can cause a problem related to the local optimal point. Thus, stochastic mutation needs to be performed on a newly generated chromosome. Finally, the global optimal genes for a given problem remain. This process is repeated until the predetermined requirements are met [50,51].

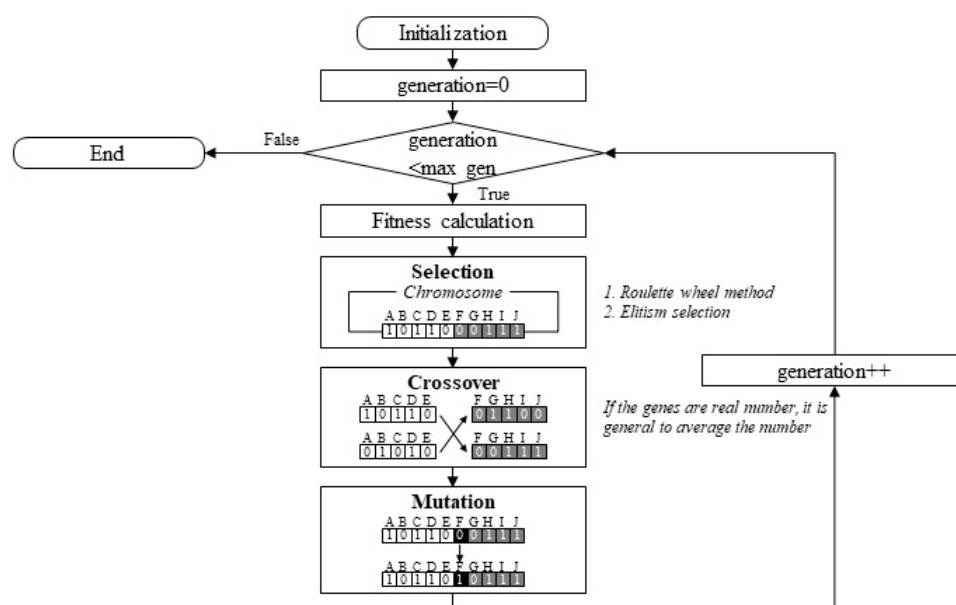


Figure 1. Genetic algorithm process.

### 3. Model Development

As discussed in the previous sections, a model is developed for estimating the service life of MEP components based on previous data. The model comprises four different modules: (1) data collection, (2) attribute selection, (3) attribute weighting based on GA, and (4) case retrieval. In the data collection module, a case base is constructed from the maintenance-related data of past projects. The acquired cases need to be carefully reviewed because they have a major effect on the estimation accuracy [39,43,53]. For research progress, GA and WED are adopted as component methodologies.

More specifically, the process of the estimation model is initiated by inputting the extracted attributes. Then, the optimal attributes are determined by changing iteratively attribute weights based on GA. The attribute weights representing each case need to be reliably computed because the weights are essential to measuring the degree of similarity between cases. Meanwhile, the case retrieval module computes the similarity scores for the test cases. In this process, similarity function based on the WED with attribute weights is used. According to the similarity scores, similar cases are extracted from database. The estimated and original service lives are compared to validate the developed model. In the following sections, the specific processes of the developed model illustrated in Figure 2 are explained.

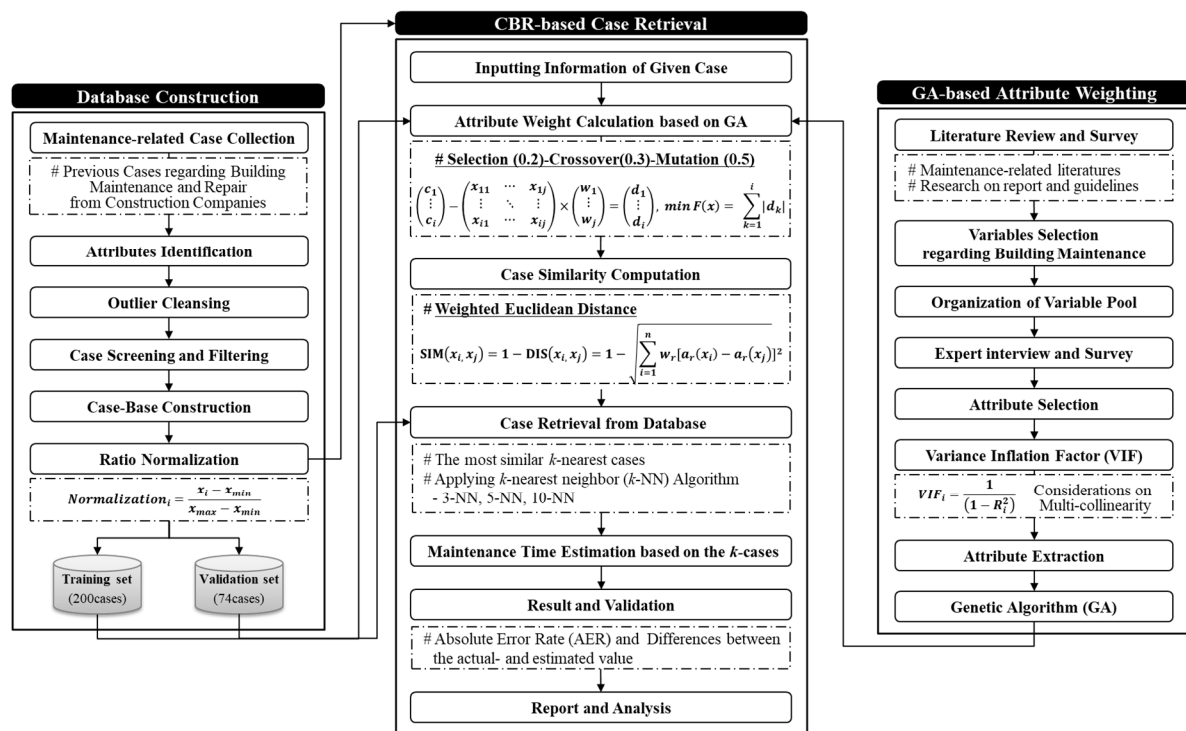


Figure 2. Case-based reasoning (CBR)-based service life estimation model.

#### 3.1. Data Collection

In order to examine the reliability and usefulness of the proposed model, a database is established based on previous data provided by Korea Land and Housing Corporation (LH). This is a recognized public corporation that has conducted and managed a number of construction projects over a long period of time. The acquired data from LH can be reliably used to estimate the service life for building maintenance because they include various maintenance-related activities and information that occurred in the past. A model based on such data will enable more accurate and reliable estimation. This research focused on the service life of MEP components because the related repairs account for a large portion of the total building maintenance [12]. MEP components consist of various subsystems including heating, air conditioning, ventilation, lighting, power, pipelines, elevators, electricity, and water

pumps [12,13]. These components are closely associated with the quality and comfort of residents' lives in apartment buildings.

In order to construct the database for estimating the service life of the building, historical cases that include the maintenance or repairs of buildings are carefully reviewed. The cases provide diverse information, such as the management area, heating system, areas for management, completion year, age of building, maintenance cost, and number of repairs, in addition to general building information such as the building coverage ratio, floor area ratio, number of buildings, number of floors, number of households, and number of parking lots. Apartment buildings completed from 1978 to 2009 are used to construct the database.

In total, 712 cases related to the maintenance of apartment buildings were collected to establish the database. However, cases including erroneous data or omitted values may not only decrease the reliability of the database but also negatively affect the accuracy of the estimated results [43,54]. To ensure the reliability and consistency of the collected cases, outliers with inappropriate data were excluded and screened from the database [34,35]. In total, 276 of the collected cases were selected. These were preprocessed for computing the case similarity and retrieving similar cases. More specifically, the raw data in the cases need to be standardized and normalized because the attributes are expressed on various scales [31,34]. Ratio standardization is applied because it is a commonly used method to present raw data from different cases on an equivalent scale [34,38,55]. The normalization is an important process because it can reduce distortions in the similarity measurement that stem from data with different scales [38]. The raw data are normalized as follows.

$$Normalization_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

where  $x_i$  is the value of attribute  $i$  and  $x_{min}$  and  $x_{max}$  are the minimum and maximum values, respectively, of each attribute. The normalized data are used to compute the case similarity, which is discussed in the next section.

### 3.2. Attribute Selection

The process of selecting and extracting attributes is described here. In order to estimate the service life for building maintenance with CBR, the attributes affecting the repair and maintenance of apartment buildings should be selected and weighted with GA. This is because the weights of the attributes are essential to calculating the case similarity. Figure 3 illustrates the process of selecting and extracting the attributes. First, relevant variables are investigated. In order to establish a variable pool, various papers and reports in the literature addressing maintenance cost estimation methodology [3,17,18], service life pattern identification [1,12], strategic planning for maintenance [20–22], and statistics and guidelines regarding building maintenance and repair [9,56] are extensively reviewed.

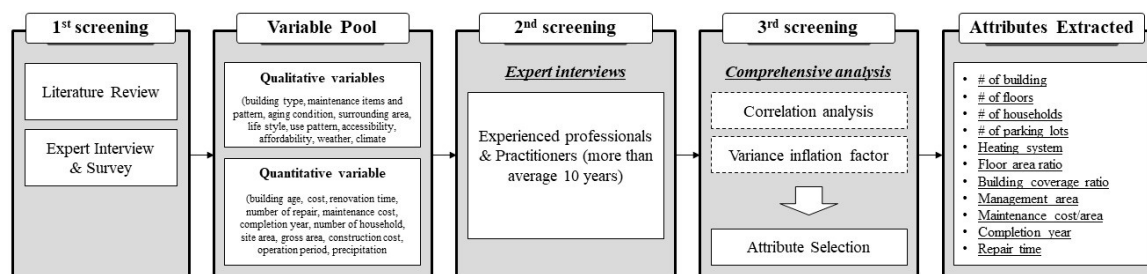


Figure 3. Attribute selection and extraction process.

Based on the variables' pool, maintenance-related factors are identified. The next step is to derive key attributes closely associated with maintenance. Twenty-three factors were extracted, then, 20 experts were interviewed to identify key attributes. These experts had experienced careers in the construction



industry covering fields such as cost management (5 persons), construction management (4 persons), safety and health management (3 persons), environment management (3 persons), and maintenance management (5 persons). They are thought to be qualified to determine the key attributes. Based on the interviews, 11 attributes were selected for measuring the similarity and estimating the service life of MEP components. The selected attributes were divided into project-related general data and maintenance-related data. The maintenance-related data included the management area (MA), heating system (HS), completion year (CY), maintenance cost per unit area (MC), and service life (SL) because they are closely associated with the deterioration of a building. In addition, project-related data such as the number of buildings (NB), number of floors (NF), number of households (NH), parking lots (PL), floor–area ratio (FA), and building–coverage ratio (BC) were included to search for previous cases similar to a given case [43]. Furthermore, the multicollinearity among the extracted attributes was examined because the attributes comprising a case have a high possibility of covariance, which can reduce the accuracy and reliability of the estimated results [31,53,57]. The variance inflation factor (VIF) was used to identify correlations among attributes. The degree of multicollinearity is calculated as follows.

$$VIF_i = \frac{1}{(1 - R_i^2)} \quad (2)$$

where  $R_i^2$  is the coefficient value obtained by regression of all independent attributes in the dataset. The VIF scores of NB, NF, NH, PL, FA, BC, HS, MA, CY, and MC were 2.004, 3.533, 1.507, 3.259, 2.274, 1.188, 1.180, 1.495, 1.736, and 1.219, respectively. Attributes with a VIF score of 5 were not observed, which indicates insignificant multicollinearity between the attributes [43,58,59]. Thus, the final 10 input attributes were extracted: NB, NF, NH, PL, FA, BC, HS, MA, CY, and MC. In addition, SL was set as the output attribute. The remaining attributes that are excluded from the inputs and output were sorted as reference information.

### 3.3. Attribute Weighting Based on GA

As discussed previously, this research adopts a genetic algorithm (GA) to weight the extracted attributes. This is an optimization algorithm that simulates an evolutionary process [38,45,60]. As shown in Figure 1, GA consists of four steps: initialization, fitness calculation, selection and crossover, and mutation.

1. **Initialization.** In this research, values between 0 and 1 are used as a chromosome (i.e., a set of attribute weights). Randomly selected weights are utilized to initialize the optimization process. A chromosome is made up of nine genes representing the attribute weights.
2. **Fitness calculation.** For an efficient search, it is necessary to select an appropriate fitness function for determining superior chromosomes. In this research, the error ratio of the training set was used as a fitness function to determine superior chromosomes.
3. **Selection and Crossover.** Parent chromosomes with the highest fitness are generally selected to generate offspring. However, this method is not suitable for global optimization because it greatly degrades the diversity. In this study, the roulette wheel and elitism selection methods are employed to find superior chromosomes. Crossover is applied to the selected two chromosomes.
4. **Mutation.** To avoid the problem of falling into a local optimal point, random mutation of a newly generated chromosome is necessary. Thus, the weights of some attributes in one chromosome are randomly changed.

In summary, GA is used to weight the attributes. One chromosome is produced based on the following ratio; roulette wheel (0.1), elitism (0.1), crossover (0.3), and mutation (0.5). The process is repeated to derive the optimal weights. Of the 276 total cases, 201 cases are used as a training set to find the optimal weight. The remaining cases are used to validate the developed model in the experiment. Table 3 indicates that CY (0.3530) and MC (0.3529) are critical attributes. This may be because these attributes are closely related to the deterioration of building maintenance.

**Table 3.** Configuration of extracted attributes and weight based on genetic algorithm (GA).

Data Type	Attribute	Attribute Type	Measurement Scale	Attribute Weight
Project-related general data	Building coverage ratio (BC)	Numeric	Percentage (%)	0.03545
	Floor area ratio (FA)	Numeric	Percentage (%)	0.03629
	Number of building (NB)	Numeric	Integer	0.04869
	Number of floor (NF)	Numeric	Integer	0.02921
	Number of households (NH)	Numeric	Integer	0.03283
Maintenance-related data	Parking lots per household (PL)	Numeric	Ratio	0.03267
	Heating system (HS)	One of a list	Binary (1 or 0)	0.03942
	Management Area (MA)	Numeric	Real number (m <sup>2</sup> )	0.03946
	Completion year (CY)	Numeric	Real number (year)	0.35300
Output	Maintenance cost per unit area (MC)	Numeric	Real number	0.35298
	Service life (SL)	Numeric	Real number (year)	-
Reference data	Number of maintenance (NM)	Numeric	Integer	-
	Total maintenance cost	Numeric	Real number	-

### 3.4. Case Retrieval

The cases in the database can be represented as attributes and solutions [45]. The case similarity can be computed based on the attribute weights and similarity function. As described previously, GA is used as a method for assigning weights for the selected attributes. Then, the case similarity is computed based on data normalized with the NORMSDIST function of Microsoft Excel.

This is essential to the case retrieval phase of CBR [23,39,41,43,61]. This is because previous cases are extracted depending on the similarity score, and an appropriate similarity measurement can reveal hidden relationships between attributes comprising cases [38,61]. Various similarity measurements are available, such as the WED, Minkowski distance, Mahalanobis distance, and arithmetic summation [31,53,62,63]. This research adopts WED to calculate the similarity because it is commonly used [34,35,37,43]. In Euclidean space, the similarity is computed based on the relative locations of attributes, which represent each case in the database [38,43]. The similarity distance between cases is determined by the square root of the sum of the squares of the differences between attributes [34,37]:

$$\text{SIM}(x_i, x_j) = 1 - \text{DIS}(x_i, x_j) = 1 - \sqrt{\sum_{r=1}^n w_r [a_r(x_i) - a_r(x_j)]^2} \quad (3)$$

where  $\text{SIM}(x_i, x_j)$  is the similarity value for the independent cases  $x_i$  and  $x_j$  and  $\text{DIS}(x_i, x_j)$  is the relative distance between these cases. In addition,  $a_r(x)$  is the value of the  $r^{\text{th}}$  attribute of case  $x_i$  and  $w_r$  is the weight of the case determined from GA [31,34,38,43]. This equation is used to determine similarities between selected cases. The most similar cases are retrieved according to the similarity scores. The k-nearest neighbor (NN) algorithm is used to estimate the service life based on retrieved cases, where the estimated values are the average of  $k$  similar cases [27,34,43]. In the estimation, the use of redundant nearest neighbors may cause the extraction of dissimilar cases, which can produce inaccurate results [43,49]. On the other hand, a limited number of nearest neighbors will also reduce the reliability and accuracy of the estimated results. Determining the optimal nearest neighbors is a concern in CBR. In this research, the service life is estimated based on outputs determined from 3, 5, and 10-NN.

## 4. Experiment

### 4.1. Experimental Process

This research focused on estimating the service life for maintenance based on previous cases. To verify the applicability of the developed model, an experiment was performed as schematized in Figure 4. First, 20 test cases were randomly selected. Second, GA was employed to weight attributes based on 201 cases in the training dataset. The weights were used to compute the similarity distance between the test cases and collected cases. The remaining 75 cases in the validation dataset were used

to retrieve cases similar to the test case. Based on the similarity scores,  $k$ -similar cases were extracted from the database. In this research, ten similar cases were retrieved for each test case. The retrieved cases were used as the essential data to estimate the service life of the MEP components. Finally, the estimated values were validated by a comparison to the actual service life in terms of the absolute error ratio (AER) and difference.

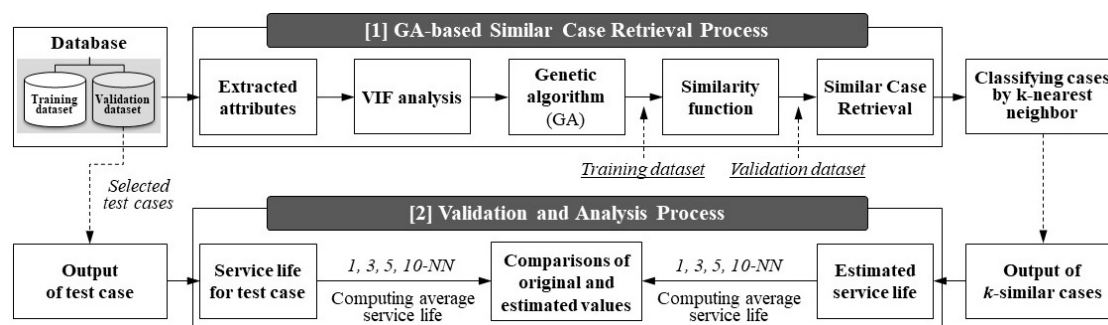


Figure 4. Experiment Process.

#### 4.2. Results and Discussion

As noted previously, 201 training cases were used to determine the attribute weights in the experiment. As indicated in Table 3, the weights were calculated in accordance with GA process in Java. The error rates were calculated for the training set based on leave-one-out validation (LOOCV), which is a type of  $k$ -fold cross-validation [31]. The error rates of 1-, 3-, 5-, and 10-NN were 8.84, 7.17, 6.96, and 7.43%, respectively (see Table 4). This indicates that the derived outcomes are reliable and accurate with an error ratio of about 7%. However, such results were determined based on the optimized weights for the training set. This implies that inaccurate results may be produced if they are applied to different datasets. Thus, a further experiment was performed on the validation dataset to confirm the applicability of the developed model.

Table 4. Case similarity and error rate for 201 training sets

Classification	$k$ -Nearest Neighbors			
	1-NN	3-NN	5-NN	10-NN
Similarity	98.88%	98.66%	98.52%	98.28%
MAER (%)	8.84%	7.17%	6.96%	7.43%

An experiment was carried out on 75 cases in the validation dataset based on the 20 randomly selected test cases. First, similarity scores were calculated based on WED. Then, cases similar to each test case were extracted. The estimated and real service lives were compared with 1-, 3-, 5-, and 10-NN. Table 5 presents the profile and input attributes that were randomly selected for model validation.

Table 6 presents the similarity scores summarized by the 1-, 3-, 5-, and 10-NN approaches. The similarities ranged from 85.2 to 99.8%, which are high similarity scores. The overall similarity and standard deviation (SD) of 10-NN were 97.4% and 0.033, respectively, which indicates that most cases extracted by 10-NN are similar to the given cases. However, some cases such as T8 showed a lower similarity score. The similarities of 1-, 3-, 5-, and 10-NN for T8 were 88.7, 86.7, 85.9, and 85.2%, respectively. This seems to be because the cases similar to the T8 are lack in the database. The similarity scores were used to extract cases and estimate the service life for building maintenance [31,34,43].

Table 7 presents the experimental results for the actual and estimated service lives of the MEP components. The service lives estimated with 1-, 3-, 5-, and 10-NN were compared with the actual service life according to the AER and difference. The mean absolute error rates (MAERs) with 1-, 3-, 5-, and 10-NN were 8.19%, 6.92%, 6.63%, and 7.48%, respectively. The maximum and minimum

AERs were 27.27% and 0%, respectively, for 1-NN; 24.24% and 0%, respectively, for 3-NN; 25.45% and 1.05%, respectively, for 5-NN; and 26.25% and 0.80%, respectively, for 10-NN. Overall, the estimated values showed reliable accuracy with an error rate of about 7%. However, the error rates with 1-, 3-, 5-, and 10-NN were high for several test cases: 12.50%, 20.83%, 15.00%, and 26.25%, respectively, for T2; 23.81%, 19.05%, 18.10%, and 17.14%, respectively, for T8; and 27.27%, 24.24%, 25.45%, and 17.27%, respectively, for T14. This may be because cases with low similarity were retrieved. However, this may not necessarily be due to low similarity scores [43]. The estimation results with CBR are case-sensitive, which implies that the estimated results can change depending on the output of the retrieved cases [31,34]. Therefore, even though the similarities between the test and retrieved cases were very high, the estimated output can differ from the actual value of the test case depending on the project conditions or unique characteristics. In this regard, the lower performance can be explained because some extreme cases with disparate service lives were included in the retrieved cases. According to Christensen and Dysert [64], the general range for low variation is from −20% to −50%, and the range for upper variation is from +30% to +100%. Therefore, the slightly high AER in some cases may be even acceptable for estimation during the preconstruction phase for which relevant data are insufficient [43,64]. Therefore, this present that the developed model is applicable to estimating the service life of building components. The model can be used by building managers in the preconstruction phase to identify when major maintenance-related problems will occur.

**Table 5.** Profiles and input attributes for the test cases.

Case Number	Project-Related Attributes						Maintenance-Related Attributes			
	BC	FA	NB	NF	NH	PL	HS	MA	CY	MC
T1	28%	374%	2	15	299	1.5	1	27,604 m <sup>2</sup>	1992	0.0132
T2	21%	360%	3	25	214	1.0	0	21,278 m <sup>2</sup>	2003	0.2430
T3	18%	296%	2	23	408	1.0	0	42,987 m <sup>2</sup>	1998	0.4533
T4	30%	88%	28	3	762	0.4	0	63,412 m <sup>2</sup>	1988	1.2606
T5	20%	204%	28	14	4424	1.1	1	471,336 m <sup>2</sup>	1979	0.3033
T6	26%	302%	2	22	329	1.1	0	35,088 m <sup>2</sup>	1996	0.1618
T7	22%	281%	2	18	182	0.9	0	21,445 m <sup>2</sup>	1993	6.7615
T8	17%	192%	25	15	3481	0.5	1	2,374,050 m <sup>2</sup>	1990	0.3855
T9	19%	218%	7	15	468	1.5	0	67,027 m <sup>2</sup>	1993	4.5238
T10	27%	385%	5	26	791	0.9	0	87,872 m <sup>2</sup>	1996	0.0541
T11	15%	208%	10	15	1800	0.4	0	106,560 m <sup>2</sup>	1992	0.4603
T12	20%	200%	2	13	214	1.0	1	21,602 m <sup>2</sup>	1988	0.3045
T13	14%	179%	4	15	375	1.0	1	95,337 m <sup>2</sup>	1984	1.2636
T14	21%	123%	6	6	300	1.0	0	47,430 m <sup>2</sup>	2003	0.1951
T15	14%	115%	9	9	432	0.7	0	41,893 m <sup>2</sup>	1985	2.6308
T16	21%	198%	10	15	700	1.1	0	62,776 m <sup>2</sup>	1988	0.0717
T17	17%	232%	12	15	1980	0.2	1	169,224 m <sup>2</sup>	1988	4.1010
T18	22%	293%	5	13	312	0.5	0	74,751 m <sup>2</sup>	1979	0.2332
T19	21%	283%	5	25	345	1.2	0	61,388 m <sup>2</sup>	1999	1.3847
T20	12%	138%	122	24	5540	1.0	1	646,257 m <sup>2</sup>	1988	2.6194

Note: BC = Building coverage ratio; FA = Floor area ratio; NB = Number of building; NF = Number of floor; NH = Number of households; PL = Parking lots per household; HS = Heating system; MA = Management Area; CY = Completion year; MC = Maintenance cost per unit area.

**Table 6.** Case similarity for 20 test cases.

NN	Number of Test Cases																			
	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20
1	98.5	99.1	99.4	97.7	94.4	99.7	99.7	88.7	99.0	99.1	98.5	99.3	99.6	98.6	98.9	99.5	98.7	99.8	93.0	99.0
3	98.1	99.1	99.4	97.6	94.3	99.4	99.5	86.7	99.0	99.0	98.4	98.9	99.5	98.4	98.9	99.4	98.7	99.0	92.9	98.7
5	98.0	99.0	99.3	97.6	94.2	99.3	99.4	85.9	98.9	98.9	98.3	98.7	99.4	98.2	98.9	99.3	98.7	98.8	92.8	98.7
10	97.7	98.8	99.1	97.5	94.1	99.1	99.3	85.2	98.9	98.7	98.2	98.4	99.1	97.9	98.6	99.2	98.4	98.3	92.5	98.6

Note: Unit: percentage (%)

Table 7. Experiment results.

Case Number	Service Life of MEP Component					Difference				Absolute Error Rate (AER, %)			
	Actual (Years)	Estimated (Years)											
		1-NN	3-NN	5-NN	10-NN	1-NN	3-NN	5-NN	10-NN	1-NN	3-NN	5-NN	10-NN
T1	19	20	20.0	18.8	18.7	−1.0	−1.0	0.2	0.3	5.26%	5.26%	1.05%	1.58%
T2	8	9	9.7	9.2	10.1	−1.0	−1.7	−1.2	−2.1	12.50%	20.83%	15.00%	26.25%
T3	15	14	15.0	14.8	14.2	1.0	0.0	0.2	0.8	6.67%	0.00%	1.33%	5.33%
T4	25	26	25.0	24.6	24.7	−1.0	0.0	0.4	0.3	4.00%	0.00%	1.60%	1.20%
T5	34	30	31.7	29.0	27.5	4.0	2.3	5.0	6.5	11.76%	6.86%	14.71%	19.12%
T6	17	17	16.7	16.8	16.4	0.0	0.3	0.2	0.6	0.00%	1.96%	1.18%	3.53%
T7	19	20	19.0	18.8	19.2	−1.0	0.0	0.2	−0.2	5.26%	0.00%	1.05%	1.05%
T8	21	26	25.0	24.8	24.6	−5.0	−4.0	−3.8	−3.6	23.81%	19.05%	18.10%	17.14%
T9	18	19	20.0	20.0	20.1	−1.0	−2.0	−2.0	−2.1	5.56%	11.11%	11.11%	11.67%
T10	17	14	15.7	16.0	16.4	3.0	1.3	1.0	0.6	17.65%	7.84%	5.88%	3.53%
T11	20	21	21.0	20.6	20.6	−1.0	−1.0	−0.6	−0.6	5.00%	5.00%	3.00%	3.00%
T12	25	25	24.7	25.4	25.4	0.0	0.3	−0.4	−0.4	0.00%	1.33%	1.60%	1.60%
T13	29	29	28.7	28.6	28.2	0.0	0.3	0.4	0.8	0.00%	1.15%	1.38%	2.76%
T14	11	8	8.3	8.2	9.1	3.0	2.7	2.8	1.9	27.27%	24.24%	25.45%	17.27%
T15	28	25	25.3	26.2	26.6	3.0	2.7	1.8	1.4	10.71%	9.52%	6.43%	5.00%
T16	25	23	23.7	24.2	24.0	2.0	1.3	0.8	1.0	8.00%	5.33%	3.20%	4.00%
T17	25	25	24.7	24.4	24.8	0.0	0.3	0.6	0.2	0.00%	1.33%	2.40%	0.80%
T18	34	31	30.7	30.2	28.3	3.0	3.3	3.8	5.7	8.82%	9.80%	11.18%	16.76%
T19	13	12	13.3	13.2	13.5	1.0	−0.3	−0.2	−0.5	7.69%	2.56%	1.54%	3.85%
T20	26	25	24.7	24.6	24.9	1.0	1.3	1.4	1.1	3.85%	5.13%	5.38%	4.23%
Minimum absolute error rate										0.00%	0.00%	1.05%	0.80%
Mean absolute error rate (MAER)										8.19%	6.92%	6.63%	7.48%
Maximum absolute error rate										27.27%	24.24%	25.45%	26.25%

Table 8 presents reference information for T10 that can be utilized as important data for building maintenance management, such as component types for maintenance, number of maintenances, and overall maintenance cost. Managers can use this information to assume how much building maintenance will occur after completion. More specifically, the reference information shows that components such as  $M_2$  (27.87%),  $P_2$  (13.11%), and  $P_3$  (24.59%) should be carefully considered. In other words, managers can identify factors regarding which MEP components will primarily cause maintenance-related problems. Furthermore, Figure 5 illustrates the probability distributions of the service life and maintenance cost based on MCS, which is a computational approach that depends on random samplings to gain statistical results [65,66]. These distributions were determined based on the beta-PERT distribution, which is a recognized method for estimating values from insufficient information because it can handle diverse types of skewness [43,67,68]. In this research, 10,000 iterations were carried out to determine the distributions of the components.

Table 8. Reference information of retrieved cases for T10.

NN	MEP Component											Service Life (Years)	Maintenance Cost (US\$)
	$M_1$	$M_2$	$M_3$	$M_4$	$E_1$	$E_2$	$E_3$	$E_4$	$P_1$	$P_2$	$P_3$		
1	0	2	0	0	0	1	0	1	0	0	1	25	41,992
2	0	0	0	0	0	0	0	0	0	0	3	16	4,584
3	0	0	0	0	0	0	1	0	0	1	1	10	24,321
4	0	6	0	7	1	0	0	0	0	6	6	7	106,528
5	3	0	0	0	0	0	0	0	0	0	1	10	241,796
6	0	7	0	0	0	0	0	1	0	0	1	11	16,663
7	1	0	0	0	0	0	0	0	0	0	0	10	85,853
8	0	2	0	0	0	0	0	0	0	1	1	30	59,236
9	0	0	0	0	0	2	0	0	0	0	1	10	130,198
10	2	0	0	0	0	0	1	0	0	0	0	12	14,380
Avg.	0.6	1.7	0	0.7	0.1	0.3	0.2	0.2	0	0.8	1.5	14.1	72,555
%	9.84%	27.87%	0.00%	11.48%	1.64%	4.92%	3.28%	3.28%	0.00%	13.11%	24.59%	-	-

Note:  $M_1$  = Heating facility;  $M_2$  = Elevator and lift;  $M_3$  = Ventilation facility;  $M_4$  = Security/crime prevention facility;  $E_1$  = Spare power facility;  $E_2$  = Substation;  $E_3$  = Lightning protection facility and outdoor lighting;  $E_4$  = Communication and broadcast;  $P_1$  = Hot-water supply facility;  $P_2$  = Water pump;  $P_3$  = Drainage facility. Maintenance cost is converted from South Korea won (KRW) to U.S. dollars (USD); exchange rate: KRW/USD = 1134.5 (prevailing rate on 1 April 2019).



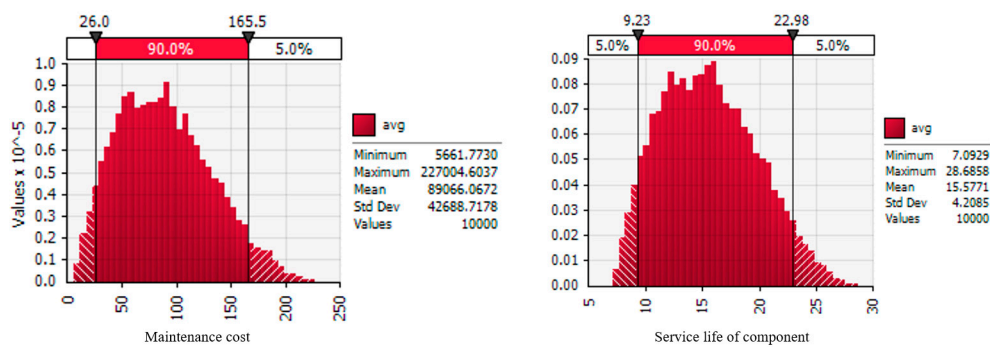


Figure 5. Service life distribution for T10 (example).

The distributions indicate that the minimum, average, and maximum maintenance costs were 5661, 89,066, and 227,004 USD, respectively, and the minimum, average, and maximum service lives were 7.09, 15.57, and 28.68 years, respectively. By synthesizing the information with the estimated results, building managers can identify not only the specific service life of building components but also the number of repairs and approximate maintenance cost in a comprehensive manner. Therefore, the developed estimation model can support more systematic building maintenance than current approaches because the outcomes are determined based on reliable data extracted from previous cases. It also provides various useful information for a preventive approach towards building maintenance.

## 5. Conclusions

The number of maintenance-related problems continues to increase as aging buildings deteriorate, which may cause safety-related problems and increased maintenance costs. Thus, building maintenance is globally recognized as an issue of concern, especially in densely populated areas. In particular, MEP components account for a large portion of building maintenance because they are essential to ensuring a sustainable environment and lifestyle for residents. Thus, maintenance with regard to MEP components should be carefully carried out. As a first step, the time or period for maintenance needs to be estimated. Although various studies have focused on estimating the service life for maintenance, doing so during the planning phase is very challenging because maintenance-related data and a systematic estimation methodology have not been established. To address these limitations, a model was developed to estimate the service life for building components according to previous similar cases which utilized GA and CBR. To validate the developed model, experiments were performed with 20 test cases, and the estimated results of 1-, 3-, 5-, and 10-NN were compared with the actual service life in term of AER. The experimental results indicated that the MAERs with 1-, 3-, 5-, and 10-NN were 8.19%, 6.92%, 6.63%, and 7.48%, respectively. These results demonstrate that the developed model can be used to estimate the service life of MEP components with a high reliability. The model can be used by maintenance managers to identify and evaluate when maintenance-related problems will occur. Also, it can be utilized to develop preventive measures against building deterioration and establish a maintenance strategy. The developed estimation model will help enhance existing maintenance management of buildings. This research contributes to the literature on building maintenance by proposing a method of estimating the service life of building components. In order to improve the usability of the model, the more specific service lives of individual components need to be addressed, and the number of repairs for specific components and the maintenance cost of each component should also be considered. However, the collected data did not include such specific details. This seems to be because a building consists of very diverse and complex components, which makes it difficult to record or collect specific data and information related to building maintenance from a long-term perspective. Because of these limitations, the total number of repairs and maintenance cost were presented here as reference information. Future research should address more specific content such as the maintenance cost for each trade or components that are essential to building maintenance. Furthermore, this research was primarily carried out based on CBR, which depends on the collected cases. Thus, more cases

need to be collected to improve the reliability and accuracy of the estimated results. Other attribute weighting methods should also be considered to ensure the applicability of the model. This study should improve current building maintenance and support managers in decision-making related to building repair by estimation of the service life of components.

**Author Contributions:** Conceptualization, N.K., M.P., Y.J., I.Y., and Y.A.; Investigation, K.S. and Y.J.; Methodology, N.K., K.S., I.Y., and Y.A.; Supervision, M.P., Y.J., and Y.A.; Validation, N.K., K.S., and I.Y.; Visualization, I.Y.; Writing—Original Draft, N.K., K.S., I.Y., and Y.A.; Writing—Review & Editing, N.K., M.P., and Y.A.

**Funding:** This research was funded by the Ministry of Science, ICT & Future Planning (No. 2015R1A5A1037548).

**Acknowledgments:** This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF). The authors wish to express their gratitude for the support.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Park, S.; Ahn, Y.; Lee, S. Analyzing the Finishing Works Service Life Pattern of Public Housing in South Korea by Probabilistic Approach. *Sustainability* **2018**, *10*, 4469. [\[CrossRef\]](#)
2. Chanter, B.; Swallow, P. *Building Maintenance Management*, 2nd ed.; Wiley Blackwell Publishing: Hoboken, NJ, USA, 2008.
3. Kim, S.; Lee, S.; Ahn, Y. Evaluating Housing Maintenance Costs with Loss-Distribution Approach in South Korean Apartment Housing. *J. Manag. Eng.* **2019**, *35*. [\[CrossRef\]](#)
4. Kim, J.; Han, S.; Hyun, C. Minimizing fluctuation of the maintenance, repair, and rehabilitation cost profile of a building. *J. Perform. Constr. Facil.* **2016**, *30*. [\[CrossRef\]](#)
5. Yang, I.T.; Hsieh, Y.M.; Kung, L.O. Parallel computing platform for multi-objective simulation optimization of bridge maintenance planning. *J. Constr. Eng. Manag.* **2012**, *138*. [\[CrossRef\]](#)
6. Ali, A.S.; Kamaruzzaman, S.N.; Sulaiman, R.; Peng, Y.C. Factors affecting housing maintenance cost in Malaysia. *J. Facil. Manag.* **2010**, *8*, 285–298. [\[CrossRef\]](#)
7. Kim, S.; Frangopol, D.M.; Soliman, M. Generalized probabilistic framework for optimum inspection and maintenance planning. *J. Struct. Eng.* **2013**, *139*, 435–447. [\[CrossRef\]](#)
8. Eom, C.S.; Paek, J.H. Risk index model for minimizing environmental disputes in construction. *J. Constr. Eng. Manag.* **2009**, *135*, 34–41. [\[CrossRef\]](#)
9. Ministry of Land, Infrastructure, and Transport (MOLIT). *Statistics Annual Rep*; MOLIT: Sejong, Korea, 2019.
10. Park, M.; Kwon, N.; Lee, J.; Lee, S.; Ahn, Y. Probabilistic Maintenance Cost Analysis for Aged Multi-Family Housing. *Sustainability* **2019**, *11*, 1843. [\[CrossRef\]](#)
11. Shohet, I.M.; Puterman, M.; Gilboa, E. Deterioration patterns of building cladding components for maintenance management. *Constr. Manag. Econ.* **2002**, *20*, 305–314. [\[CrossRef\]](#)
12. Lee, S.; Ahn, Y. Analyzing the Long-Term Service Life of MEP Using the Probabilistic Approach in Residential Buildings. *Sustainability* **2018**, *10*, 3803. [\[CrossRef\]](#)
13. Ministry of Land, Infrastructure, and Transport (MOLIT). *Korean Apartment Housing Management Act*; MOLIT: Sejong, South Korea, 2016.
14. Khanzode, A.; Fischer, M.; Reed, D. Benefits and lessons learned of implementing building virtual design and construction (VDC) technologies for coordination of mechanical, electrical, and plumbing (MEP) system on a large healthcare project. *J. Inf. Technol. Constr.* **2008**, *13*, 324–342.
15. Ho, D.C.; Gao, W. Collective action in apartment building management in Hong Kong. *Habitat Int.* **2013**, *38*, 10–17. [\[CrossRef\]](#)
16. Mohamad, S.B.H.S.; Rhaman, M.A.A.; Akasah, Z.A. A review of the maintenance performance factors for heritage buildings. In Proceedings of the International Civil and Infrastructure Engineering Conference, Kota Kinabalu, Malaysia, 28 September–1 October 2014; Springer: Singapore, 2014; pp. 177–187.
17. Sharma, J.; Najafi, M.; Qasim, S. Preliminary cost estimation models for construction, operation, and maintenance of water treatment plants. *J. Infrastruct. Syst.* **2013**, *19*, 451–464. [\[CrossRef\]](#)
18. Stenbeck, T. Budgeting performance-based winter maintenance: Snow influence on highway maintenance cost. *J. Infrastruct. Syst.* **2009**, *15*, 251–260. [\[CrossRef\]](#)

19. Choi, K.; Kim, Y.H.; Bae, J.; Lee, H.W. Determining Future Maintenance Costs of Low-Volume Highway Rehabilitation Projects for Incorporation into Life-Cycle Cost Analysis. *J. Comput. Civ. Eng.* **2016**, *30*, 04015055. [[CrossRef](#)]
20. Arif, F.; Bayraktar, M.E.; Chowdhury, A.G. Decision support framework for infrastructure maintenance investment decision making. *J. Manag. Eng.* **2015**, *32*, 04015030. [[CrossRef](#)]
21. Cho, K.; Yoon, Y. Decision Support Model for Determining Cost-Effective Renovation Time. *J. Constr. Eng. Manag.* **2015**, *32*, 04015051. [[CrossRef](#)]
22. Elcheikh, M.; Burrow, M.P.N. Uncertainties in forecasting maintenance costs for asset management: Application to an aging canal system. *J. Risk Uncertainty Eng. Syst.* **2017**, *3*, 04016014. [[CrossRef](#)]
23. Aamodt, A.; Plaza, E. Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Commun.* **1994**, *7*, 39–59.
24. Kolodner, J. *Case-Based Reasoning*; Morgan Kaufmann: San Mateo, CA, USA, 2014.
25. Kwon, N.; Park, M.; Lee, H.S.; Ahn, J.; Shin, M. Construction Noise Management Using Active Noise Control Techniques. *J. Constr. Eng. Manag.* **2016**, *142*, 04016014. [[CrossRef](#)]
26. Riesbeck, C.K.; Schank, R.S. *Inside Case-Based Reasoning*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1989.
27. Watson, I.; Marir, F. Case-based reasoning: A review. *Knowl. Eng. Rev.* **1994**, *9*, 327–354. [[CrossRef](#)]
28. Arditi, D.; Tokdemir, O.B. Comparison of case-based reasoning and artificial neural networks. *J. Comput. Civ. Eng.* **1999**, *13*, 162–169. [[CrossRef](#)]
29. Goh, Y.M.; Chua, D. Case-based reasoning for construction hazard identification: Case representation and retrieval. *J. Constr. Eng. Manag.* **2009**, *135*, 1181–1189. [[CrossRef](#)]
30. Morcous, G.; Rivard, H.; Hanna, A.M. Case-based reasoning system for modeling infrastructure deterioration. *J. Comput. Civ. Eng.* **2002**, *16*, 104–114. [[CrossRef](#)]
31. Kwon, N.; Lee, J.; Park, M.; Yoon, I.; Ahn, Y. Performance Evaluation of Distance Measurement Methods for Construction Noise Prediction Using Case-Based Reasoning. *Sustainability* **2019**, *11*, 871. [[CrossRef](#)]
32. Kim, B.; Hong, T. Revised case-based reasoning model development based on multiple regression analysis for railroad bridge construction. *J. Constr. Eng. Manag.* **2011**, 154–162. [[CrossRef](#)]
33. Pereira, E.; Hermann, U.; Han, S.; AbouRizk, S. Case-based reasoning approach for assessing safety performance using safety-related measures. *J. Constr. Eng. Manag.* **2018**, *144*, 04018088. [[CrossRef](#)]
34. Kwon, N.; Park, M.; Lee, H.S.; Ahn, J.; Kim, S. Construction Noise Prediction Model Based on Case-Based Reasoning in the Preconstruction Phase. *J. Constr. Eng. Manag.* **2017**, *143*, 04017008. [[CrossRef](#)]
35. Ahn, J.; Ji, S.H.; Park, M.; Lee, H.S.; Kim, S.; Suh, S.W. The attribute impact concept: Applications in case-based reasoning and parametric cost estimation. *Autom. Constr.* **2014**, *43*, 195–203. [[CrossRef](#)]
36. Watson, I. *Applying Case-Based Reasoning: Techniques for Enterprise System*; Morgan Kaufmann: San Francisco, CA, USA, 1997.
37. Pal, S.K.; Shiu, S.C. *Foundations of Soft Case-Based Reasoning*; John Wiley & Sons: Hoboken, NJ, USA, 2004.
38. Ji, S.H.; Park, M.; Lee, H.S.; Ahn, J.; Kim, N.; Son, B. Military facility cost estimation system using case-based reasoning in Korea. *J. Comput. Civ. Eng.* **2011**, 218–231. [[CrossRef](#)]
39. Doğan, S.Z.; Arditi, D.; Günaydin, H.M. Determining Attribute Weights in a CBR Model for Early Cost Prediction of Structural Systems. *J. Constr. Eng. Manag.* **2006**, *132*, 1092–1098. [[CrossRef](#)]
40. Doğan, S.Z.; Arditi, D.; Günaydin, H.M. Using decision trees for determining attribute weights in a case-based model of early cost prediction. *J. Constr. Eng. Manag.* **2008**, *134*, 146–152. [[CrossRef](#)]
41. Ji, S.H.; Park, M.; Lee, H.S. Case adaptation method of case-based reasoning for construction cost estimation in Korea. *J. Constr. Eng. Manag.* **2012**, *138*, 43–52. [[CrossRef](#)]
42. Jin, R.; Cho, K.; Hyun, C.; Son, M. MRA-based revised CBR model for cost prediction in the early stage of construction projects. *Expert Syst. Appl.* **2012**, *39*, 5214–5222. [[CrossRef](#)]
43. Kwon, N.; Cho, J.; Lee, H.S.; Yoon, I.; Park, M. Compensation Cost Estimation Model for Construction Noise Claims Using Case-Based Reasoning. *J. Constr. Eng. Manag.* **2019**. accepted.
44. Koo, C.; Hong, T.; Hyun, C.; Koo, K. A CBR-based hybrid model for predicting a construction duration and cost based on project characteristics in multi-family housing projects. *Can. J. Civ. Eng.* **2010**, *37*, 739–752. [[CrossRef](#)]
45. Kim, J.K.; Kim, K. Preliminary Cost Estimation Model Using Case-Based Reasoning and Genetic Algorithms. *J. Comput. Civ. Eng.* **2010**, *24*, 499–505. [[CrossRef](#)]

46. Albright, S.C.; Winston, W.L. *Management Science Modeling*; Thomson South-Western College Publishing: Mason, OH, USA, 2007.
47. Gen, M.; Cheng, R. *Genetic Algorithms and Engineering Design*; Wiley: New York, NY, USA, 1997.
48. RazaviAlavi, S.; AbouRizk, S. Site layout and construction plan optimization using an integrated genetic algorithm simulation framework. *J. Comput. Civ. Eng.* **2017**, *31*, 04017011. [[CrossRef](#)]
49. Kim, S. Dual-Optimization Method for Improving Accuracy in GA-CBR Cost Estimating Model. Ph.D. Thesis, Seoul National University, Seoul, Korea, February 2017.
50. Mitchell, M. *An Introduction to Genetic Algorithm*; MIT Press: Cambridge, MA, USA, 1999.
51. Feng, C.W.; Liu, L.; Burns, S.A. Using genetic algorithms to solve construction time-cost trade-off problems. *J. Comput. Civ. Eng.* **1997**, *11*, 184–189. [[CrossRef](#)]
52. Goldberg, D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning*; Addison Wesley: Boston, MA, USA, 1989.
53. Ahn, J.; Park, M.; Lee, H.S.; Ahn, S.J.; Ji, S.H.; Song, K. Covariance effect analysis of similarity measurement methods for early construction cost estimation using case-based reasoning. *Autom. Constr.* **2017**, *81*, 254–266. [[CrossRef](#)]
54. Han, J.; Kamber, M.; Pei, J. *Data Mining: Concept and Techniques*; Morgan Kaufmann: Waltham, MA, USA, 2011.
55. Bod, R.; Hay, J.; Jannedy, S. *Probabilistic Linguistics*; MIT Press: Cambridge, MA, USA, 2003.
56. Marquez, A.C. *The Maintenance Management Framework, Methods and Models for Complex Systems Maintenance*; Springer: London, UK, 2007.
57. Du, J.; Bormann, J. Improved similarity measure in case-based reasoning with global sensitivity analysis: An example of construction quantity estimating. *J. Comput. Civ. Eng.* **2014**, *28*, 04014020. [[CrossRef](#)]
58. Rogerson, P.A. *Statistical Methods for Geography*; SAGE Publications Inc.: London, UK, 2001.
59. Vu, D.H.; Muttaqi, K.M.; Agalgaonkar, A.P. A variance inflation factor and backward elimination based robust regression model for forecasting monthly electricity demand using climatic variables. *Appl. Energy* **2015**, *140*, 385–394. [[CrossRef](#)]
60. Fu, Y.; Shen, R. GA based CBR approach in Q&A system. *Expert Syst. Appl.* **2004**, *26*, 167–170.
61. Burkhard, H.D. Similarity and distance in case-based reasoning. *J. Fundam. Inf.* **2001**, *47*, 201–215.
62. Zhang, Y.; Ding, L.; Love, P.E.D. Planning of deep foundation construction technical specifications using improved case-based reasoning with weighted k-nearest neighbors. *J. Comput. Civ. Eng.* **2017**, *31*, 04017029. [[CrossRef](#)]
63. Salleh, S.S.; Aziz, N.A.A.; Mohamad, D.; Omar, M. Combining Mahalanobis and Jaccard Distance to Overcome Similarity Measurement Constriction on Geometrical Shapes. *IJCSI* **2012**, *9*, 124–132.
64. Christensen, P.; Dysert, L.R. *Cost Estimate Classification System—As Applied in Engineering, Procurement, and Construction for the Process Industries*; AACE International Recommended Practices 18R-97; AACE International Recommended Practices: Morgantown, WV, USA, 2011.
65. Asmar, M.E.; Hanna, A.S.; Chang, C.K. Monte Carlo simulation approach to support alliance team selection. *J. Constr. Eng. Manag.* **2009**, *135*, 1087–1095. [[CrossRef](#)]
66. Kwon, N.; Song, K.; Lee, H.S.; Kim, J.; Park, M. Construction Noise Risk Assessment Model Focusing on Construction Equipment. *J. Constr. Eng. Manag.* **2018**, *144*. [[CrossRef](#)]
67. Pawan, P.; Lorterapong, P. A Fuzzy-Based Integrated Framework for Assessing Time Contingency in Construction Projects. *J. Constr. Eng. Manag.* **2015**, *142*. [[CrossRef](#)]
68. Salling, K.B. *Risk Analysis and Monte Carlo Simulation within Transport Appraisal*; CTT-DTU, Technical University of Denmark: Lyngby, Denmark, 2007.

