PREDICTING THE COST IMPACTS OF CONSTRUCTION NON-CONFORMITIES USING CBR-AHP AND CBR-GA MODELS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

NEŞET BERKAY DOĞAN

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
CIVIL ENGINEERING

JULY 2021

Approval of the thesis:

PREDICTING THE COST IMPACTS OF CONSTRUCTION NON-CONFORMITIES USING CBR-AHP AND CBR-GA MODELS

submitted by **NEŞET BERKAY DOĞAN** in partial fulfillment of the requirements for the degree of **Master of Science** i**n Civil Engineering, Middle East Technical University** by,

Prof. Dr. Halil Kalıpçılar

Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ahmet Türer

Head of the Department, **Civil Engineering**

Assoc. Prof. Dr. Onur Behzat Tokdemir

Supervisor, **Civil Engineering, METU**

Examining Committee Members:

Prof. Dr. İrem Dikmen Toker

Civil Engineering, METU

Assoc. Prof. Dr. Onur Behzat Tokdemir

Civil Engineering, METU

Prof. Dr. Rıfat Sönmez

Civil Engineering, METU

Assist. Dr. Güzide Atasoy Özcan

Civil Engineering, METU

Assist. Dr. Gözde Bilgin

Civil Engineering, Başkent University

Date: 13.07.2021

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name Last name : Neşet Berkay Doğan

Signature :

ABSTRACT

PREDICTING THE COST IMPACTS OF CONSTRUCTION NON-CONFORMITIES USING CBR-AHP AND CBR-GA MODELS

Doğan, Neşet Berkay
Master of Science, Civil Engineering
Supervisor: Assoc. Prof. Onur Behzat Tokdemir

July 2021, 77 pages

Quality problems in construction projects can have dramatic consequences, such as delays in timeline and cost overruns. If preventative measures are not applied, these quality problems can evolve into repetitive actions. This study introduces a proactive mechanism with the aim of protecting projects from the negative impacts of such non-conformities. It advances in three stages. The first involves determination of the attributes that comprehensively define quality problems using a literature survey and the Delphi method. The second determines the attribute weights by employing Analytical Hierarchy Process (AHP) and Genetic Algorithm (GA). In the final stage, a predictive model is developed to extract the possible outcomes of non-conformities in terms of cost impact. The predictive model adopts the case-based reasoning (CBR) approach with Mean Absolute Error (MAE) as the control criterion for prediction accuracy. Although CBR-GA yields a better MAE performance than CBR-AHP, the result is reversed for standard deviation. This thesis provides two significant outcomes in addition to the primary objective, forecasting possible failure. First, the attributes are determined to express the cases considered to contribute to the development of a record-keeping guideline for inexperienced quality practitioners. Second, the predictive model utilizes both automated and expert systems for attribute

weighting, so the study examines the effect of automated and expert systems on the model's accuracy.

Keywords: Predictive Modelling, Case-Based Reasoning, Analytic Hierarchy Process, Genetic Algorithm, Quality Problems

ÖZ

İNŞAAT PROJELERİNDEKİ KALİTE UYGUNSUZLUKLARININ MALİYET ETKİLERİNİN CBR-AHP VE CBR-GA MODELLERİYLE TAHMİN EDİLMESİ

Doğan, Neşet Berkay
Yüksek Lisans, İnşaat Mühendisliği
Tez Yöneticisi: Doç. Dr. Onur Behzat Tokdemir

Temmuz 2021, 77 sayfa

İnşaat projelerindeki kalite sorunları, zaman çizelgesinde gecikmeler ve maliyet aşımları gibi dramatik sonuçlara yol açabilir. Önleyici tedbirler uygulanmazsa, bu kalite sorunları tekrarlayan eylemlere dönüşebilir. Bu çalışma, projeleri bu tür uygunsuzlukların olumsuz etkilerinden korumak amacıyla proaktif bir mekanizmayı tanıtmaktadır. Çalışma, üç aşamadan oluşmaktadır. İlk olarak, bir literatür taraması ve Delphi yöntemi kullanılarak kalite problemlerini kapsamlı bir şekilde tanımlayan özelliklerin belirlenmiştir. Daha sonra, Analitik Hiyerarşi Süreci (AHS) ve Genetik Algoritma (GA) kullanarak öznitelik ağırlıkları belirlenmiştir. Son aşamada, maliyet etkisi açısından uygunsuzlukların olası sonuçlarını çıkarmak için tahmine dayalı bir model geliştirilmiştir. Tahmine dayalı model, Durum Tabanlı Çıkarımsama (VTÇ) yaklaşımını ve tahmin doğruluğu için kontrol kriteri olarak Ortalama Mutlak Hata (OMH) benimsemektedir. VTÇ-GA, VTÇ-AHS'den daha iyi bir OMH performansı vermesine rağmen, standart sapma için sonuç tam tersidir. Önerilen çalışma, birincil amaca ek olarak iki önemli sonuç sağlar. İlk olarak, öznitelikler, deneyimsiz kalite uygulayıcıları için bir kayıt tutma kılavuzunun geliştirilmesine katkıda bulunmaktadır. İkinci olarak, tahmine dayalı model, öznitelik ağırlıklandırma için

hem otomatikleştirilmiş hem de uzman sistemleri kullanır, bu nedenle çalışma, otomatikleştirilmiş ve uzman sistemlerin modelin doğruluğu üzerindeki etkisini de incelemektedir.

Anahtar Kelimeler: Tahmine Dayalı Modelleme, Veri Tabanlı Çıkarımsama, Analitik Hiyerarşi Süreci, Genetik Algoritma, Kalite Problemleri

To my family and my wife…

ACKNOWLEDGMENTS

I would like to express my gratitude to Asst. Prof. Dr. Onur Behzat Tokdemir for his valuable contributions to this study and my life. He has always encouraged and supported me throughout my research.

I would also thank to Bilal Umut Ayhan for his techical suggestions and comments. His constructive critisim and effort to make time despite his intensive work schedule were valuable for me.

I would like to give my special thanks to my mother, Canan Doğan, my father, Veysel Doğan, and my brother, Tevfik Orkun Doğan, who never give up believing me. They always support me at every step of my life and show their endless motivation and love to me.

I am also thankful to my close friends, Özer Gündüz, Melis Deniz, Baki Aykut Bilginer, Ergin Özgün, Sinan Özer and Erdi Komanova. They have always been there whenever I need.

Finally, I would like to thank my beloved adorable life partner, Gülin Tüfekci, for her everlasting love and emotional support. She has always been there for me with her everlasting love and supporting me in every moment in my life. I feel lucky and grateful to share my life with her.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

ABBREVIATIONS	
AHP	Analtyical Hierarchical Process
ANN	Articifial Neural Networks
ARM	Associate Rule Mining
CBR	Case-Based Reasoning
CI	Consistency Index
CR	Consistency Ratio
GA	Genetic Algorithm
IRCA	Internal Auditor Certificate or Lead Auditor Certificate
MAE	Mean Absolute Error
NCR	Non-Conformity Report
NN	Neural Networks
RC	Random Consistency Index
RMSE	Root Mean Square Error

CHAPTER 1

INTRODUCTION

Each project possesses unique characteristics, so diversity exists in the scope, type of contract, and the relevant specifications in conjunction with the challenges, encountered. These variations also lead to changes in methods of activities applied to complete a project. In other words, different techniques and approaches have been employed in construction projects, and activities are also usually non-repetitive.

Moreover, each construction project includes a level of uncertainty and risk for stakeholders, which arises the complexity of the construction projects. The challenges in uniqueness and complexity step forward the integration process to automated systems as a controlling mechanism for the construction industry. Therefore, construction projects are more prone to encounter defects and redoing the completed work or briefly rework.

Completing a project with a high-quality performance, which can be determined by comparing the completed work with specifications, is one of the most significant project success indicators. It may also prevent the contractors from delays in a schedule (Love 2002) and increase the project cost (Forcada et al. 2017; Love 2002; Love et al. 2017, 2018; Love and Sing 2013). Love (2002) stated the growth of 12,6% and 20,7% in mean cost and schedule due to rework in Australia's building projects. Moreover, the ratio of rework cost to original contract value was calculated as 0,39% (Love et al. 2018), 0,18% (Love et al. 2017), and 2,75% (Forcada et al. 2017). Besides, the contract value is in correlation with the rework cost. Since the contract values are notably high, rework leads to an enormous increase in cost and loss in profit. Love et al. (2018) indicated a 28% loss in yearly profit on average due to rework. Therefore, a remarkable profit can be made if the defects can be detected earlier.

Additionally, the defective work can evolve to repetitive action, which causes corrective and preventative measures to be more complicated and costly (Josephson and Hammarlund 1999). Therefore, there is a need for adopting a proactive quality control system that immediately notifies the practitioners about possible outcomes of defects.

In order to record the quality problems, Nonconformance Reports (NCR) are used in construction sites. Deficiencies in the quality management system may be observed by making root cause analyses of the defects reported by these NCRs. Implementation of analyses' outcomes fulfills the inadequacies. Organizational improvement can only be achieved by learning from mistakes. However, defects are not recorded or recorded from a biased perspective in construction projects (Sheng et al. 2020). It disables the problem's detection in a quality management system, benefits from the lessons learned, and even makes the adaption of an effective quality management system meaningless. One of the reasons why there is a lack of record-keeping in construction projects is that practitioners of construction projects are unfamiliar with record keeping.

This thesis' main objective is to reduce the negative impacts of quality problems in construction projects due to these problems and increase cost-performance of the projects. Following this purpose, a Case-based Reasoning (CBR)-based proactive early warning mechanism for quality management by forecasting the upcoming quality problems considering cost impact was introduced. The proposed model also enables inexperienced quality control practitioners to be informed about the drawbacks of the quality system if any precaution is not taken and provides a record-keeping system for NCRs.

CBR is a machine learning technique that benefits from historical data rather than predefined rules (Hu et al., 2016). The CBR-based models' accuracy depends on determining attributes, weights, and similarity function. Delphi Method was applied to determine the characteristics by involving the experts of quality. Then, data was modeled by converting them to the binary system with these attributes, accordingly. In this thesis, 2.527 NCR was collected from an anonymous construction company that works mainly in Russia and Eastern Europe. These NCRs were sequenced according to the occurrence time. Afterward, two techniques, Analytic Hierarchy Process (AHP) and the Genetic Algorithm (GA) were applied to specify the attribute weights. AHP is an expert system, while GA is an automated one. Therefore, comparing the results obtained from AHP and GA also enables understanding expert participation in quality-related problems. After the attributes and their weights were specified, the most similar cases to a new case were extracted to forecast the upcoming quality problem's cost impact.

This thesis was structured as follows. Chapter 2 described the literature review on quality studies. The content of the literature review fragmented regarding the type of the study, and it focused on studies that utilized a predictive model, especially adopting CBR. Chapter 3 presented the methodology of the research in detail. Attribute selection with Delphi Method, determination of attribute weights using AHP and GA, and predictive model employing CBR were introduced. The results obtained through the study were presented and discussed in Chapter 4. Finally, Chapter 5 provided a conclusion of the study and underlined significant findings and discussion as well as the limitations and future works.

CHAPTER 2

LITERATURE REVIEW

2.1. Quality Issues

Contractors are responsible for completing the projects within the project timeline and scope and with a satisfactory quality level. Quality is the most fundamental component among them since problems in quality performance significantly impact the project's implementation processes and result in irrecoverable consequences for the project's stakeholders. These consequences may be delays in the project timeline, an increase in the project cost, or harm the contractors' reputation. Therefore, many researchers have focused on quality-related issues.

Contractors expect to experience a healthy project lifecycle by completing the project on time within the scope and ultimately making a profit, which can only be achieved by decreasing the cost. However, repeating the finished work due to quality problems has a dramatic influence on cost. Thus, the cost of rework and defects were analyzed through different methods to overcome the negative effect of defects on project cost (Barber et al. 2000; Love et al. 2017, 2018; Oke and Ugoje 2013). The impact of different project types and procurement methods on the cost of rework was examined for Australian construction projects as a case study. It was found that although there is an obvious need for reducing the cost by avoiding rework, there is no significant correlation between the cost of rework and project types and procurement methods (Love 2002). Effect of project characteristics on project cost performance was also identified, and recommendations for possible solutions for the root causes of rework were provided (Hwang et al. 2009).

Additionally, Forcada et al. (2017) also specified the factors affecting rework costs, such as project characteristics and managerial issues. The direct (Love et al. 2018) and indirect (Love 2002) cost of rework were also examined, and the probabilities of both cost components were determined to derive the actual cost of the rework (Love and Sing 2013). Moreover, the impact of knowledge management (Olayinka et al. 2016) and quality program (Jafari and Love 2013) on reducing the cost of defects were investigated, and it was revealed that both had a positive impact on decreasing the cost of defects (Jafari and Love 2013; Olayinka et al. 2016).

Josephson and Hammurland (1999) questioned whether the rework cost could be reduced by early detection of defects. They concluded that even if the defects can be detected earlier, necessary actions should be applied to avoid the consequences of rework. Therefore, warning systems are required to be developed to prevent the outcomes of quality issues. Real-time monitoring systems can be an effective solution to the problem since they enable us to pursue the quality of completed work simultaneously. Zhong et al. (2018) developed a real-time monitoring system for earth-rockfill dam constructions, and Kazemian et al. (2019) were able to detect the defects of extrusion while utilizing additive manufacturing. Moreover, the quality of gravel piles can be controlled using a real-time monitoring system and the Internet of Things (Chen et al. 2020).

Although many researchers put great effort to identify the factors that lead to rework and to investigate the reasons and the outcomes of rework cost deeply, there is a gap in the literature as no proactive early warning system can notify the practitioners about the possible consequences of NCRs in terms of cost if any precaution is not taken. Above-mentioned real-time monitoring systems can detect the defects just after it occurs, and these systems enable fixing the problem and preventing higher costs. However, it cannot be warned about the issues on quality management systems by employing these systems. Hence, a system that can detect the drawbacks of the quality system and inform the practitioners about them is needed to be implemented.

2.2. Case-Based Reasoning

CBR is an AI technique that benefits from the previous knowledge learned from the previous problems' solutions. Indeed, it approaches the problems based on the principle of "similar problems have similar solutions." (Aamodt and Plaza 1994). Hence, it has been employed as an effective computerized problem-solving technique to handle the complexity of problems in the construction industry. Neural Network (NN) is an effective alternative to CBR. The accuracy obtained from NN and CBR models differs from the selected data. Although Kim et al. (Kim et al. 2005), Ozorhon et al. (Ozorhon et al. 2006), and Arditi and Tokdemir (1999a) have obtained more accurate results with CBR than NN, Kim et al. (Kim et al. 2004) have obtained more effective performance with NNs. However, there is a consensus about the flexibility of CBR models. CBR models have the advantage of accounting for the results obtained from the model (Kim et al. 2004; Ozorhon et al. 2006) and handling the missing information (Kim et al. 2005). Moreover, NN needs longer times to compute while testing the model (Ozorhon et al. 2006), and this is a drawback since responses obtained in a short time are a remarkable benefit for warning systems.

Accurate estimation of cost in the early stages of construction projects is challenging since the information is minimal at the beginning of the project. The historical data becomes significant in the cost estimation at the conceptual or planning phase of the projects. Therefore, CBR is used as a cost estimation tool for construction projects (Ahn et al. 2020; An et al. 2007; Choi et al. 2014; Jin et al. 2012, 2014; Kim et al. 2004; Kim and Kim 2010; Kim 2012; Koo et al. 2010; Lee et al. 2013) and some researches were mainly focused on the estimating the cost of railroad-bridge construction projects (Kim and Hong 2012), military facilities (Ji et al. 2011), multi-family housing projects (Hong et al. 2011; Ji et al. 2010; Koo et al. 2010) and pump station construction projects (Marzouk and Ahmed 2011). Although many pieces of research have been conducted on cost estimation for construction projects, quality issues have not been targeted in terms of cost via CBR.

Several parties such as contractors, employers, sub-contractors, and designers, etc. involve in construction projects. Since each has its interest, it is inevitable to conflict of interests of the involved parties. Mediation is a relatively less expensive solution for conflict of interest or disputes. Li (1996) has developed a CBR-based negotiation model, named MEDIATOR, to support construction parties' negotiation process. This mediation model also provides a fair and neutral approach to the negotiation. Litigation is another solution for cases that parties cannot negotiate. However, litigation is too expensive since it is difficult to find experts who have a legal background in engineering issues. It is beneficial to estimate the litigation outcome; therefore, CBR-based models have been developed to avoid litigation's unnecessary expense (Arditi and Tokdemir 1999b; Chen and Hsu 2007; Cheng et al. 2009).

Planning, risk, and international market selection are the other areas of the construction industry that have been utilized from CBR-based models. CBR-based methodologies that used previous scheduling experience were proposed to facilitate the preparation process (Dzeng and Tommelein 1997, 2004; Ryu et al. 2007). CBR has also been applied to select risk response strategies (Fan et al., 2015; Forbes et al., 2007). Moreover, Ozorhon et al. (2006) developed a model to ease the international market selection.

Determination of attribute weight is one of the most significant criteria for obtaining successful and better results in CBR. The researchers mentioned above have employed different methods for attribute weights calculation. Feature counting(Ahn et al. 2020; Doğan et al. 2006), gradient descent(Ahn et al. 2020; Doğan et al. 2006), multi-regression analysis (Jin et al. 2012), and decision trees (Doğan et al. 2008) are the methods applied for the determination of attribute weights. However, GA is the most popular approach selected for weight determination (Choi et al., 2014; Doğan et al., 2006; Hong et al., 2011; Ji et al., 2011; Kim and Kim, 2010). Moreover, it was stated by An et al. (An et al. 2007) that due to the complexity of problems in the construction industry, it is essential to involve experts during the solution process rather than relying only on the computational approaches. Therefore, both GA as a computerized process and AHP, which is an expert system, are employed in this thesis in order to investigate whether involving an expert system on quality-related problems can improve the accuracy of the proposed methods.

CHAPTER 3

METHODOLOGY

A probabilistic model for early detection of the problems in a quality system was introduced in this thesis. NCRs obtained from the construction site are expressed with the attributes. Hence, the process commences with determining the attributes via Delphi Method. After selecting the attributes with the experts, the collected data are converted into binary format according to these attributes.

CBR benefits from the experiences gained from previous similar cases. The success of the CBR model mainly depends on the selection of attributes weights. Therefore, two methods, AHP and GA, were implemented to determine the attribute weights. It was also aimed to reveal whether an expert system or an automated system is more beneficial for the CBR models applied to quality problems in construction projects.

When a new case is introduced to the model, the system was detected the most similar cases that fulfill the predetermined threshold. Ten cases followed by similar cases determined via CBR were extracted, and the probability of the cases was calculated in terms of cost impacts. Finally, the obtained results were compared with the actual data to determine the accuracy of the model.

3.1. Attribute Selection with Delphi Method and Data Preparation

The thesis adopts the Delphi method to determine the stimulating attributes for reworks in construction projects. An up-to-date and wide range of literature reviews was performed to capture the leading attributes. Then, these were ranked via the Delphi method so that it obtains a comprehensive and confined list. The Delphi method includes iterative processes that analyze a statistical group of responses and be capable of receiving reliable results. Since the responses are taken anonymously from the participants, responses cannot influence others' opinions. In other words, this method encourages participants to reflect on their opinion without feeling any pressure. The steps of Delphi can be summarized as follow:

Firstly, an adequate number of experts or panelists experienced in Quality management in the construction industry should be identified. The recommended number for panelists was defined between 10 and 20 (Ayhan and Tokdemir 2019a; Hallowell and Gambatese 2010) Therefore, we decided to contact 11 panelists to advance in Delphi. Table 3.1 describes the required qualifications that panelists should endow. The panelist constituted mechanical and civil engineers as well as architectures who are currently working in international construction companies or universities. Besides Internal Auditor Certificate or Lead Auditor Certificate (IRCA), it is compulsory for those working in a construction company. However, this criterion was ignored when the participant has vast experience in quality management. On the other hand, an academic background in construction management, specifically quality management, is required in order to be a panelist if s/he participated from the university. Therefore, 11 panelists attended concerning these criteria given in Table 3.1, and their details are demonstrated in Table 3.2, respectively.

The second step was to prepare the questionnaire for ranking the attributes. Participants gave a score between 1-7 from strongly disagree to agree strongly. There were three critical conditions for discussion. Before interpreting the results, the mean of feedbacks and standard deviations should be defined clearly. Mean values indicate the central tendency, whereas the standard deviation accounts for the degree of consensus (Kuzucuoğlu et al., 2019; Ayhan and Tokdemir 2019a; Seyis and Ergen 2017). The first discussion was that the attributes with low mean values were eliminated from the current list extracted from the literature review. The second one is valid for the high-mean score attributes with significant standard deviations. Higher standard deviations implied no consensus among the participants, as indicated above. Therefore, more than one round was carried out to satisfy the agreement between them. The third condition was the highest score with low standard deviations. The higher mean values with lower standard deviations should be provided to achieve this condition.

Table 3.1. *Required Criteria for being a panelist*

Requirement	Educational degree
Ed1	B.S. taken from department listed below: - (Ed1-1) Mechanical Engineering - (Ed1-2) Civil Engineering - (Ed1-3) Architecture
Ed2	At least one of the certificates specified below: - (Ed2-1) Auditor Certificate - (Ed2-2) Lead Auditor Certificate
Ed3	Having a graduate-level background in construction management or quality management
	Experience level
Ex1	At least ten years experience in the construction industry
Ex2	At least 5 years experience in quality control and management

Figure 3.1 exhibits the study plan for the Delphi process. The literature knowledge composed the base of attributes for reworks. A wide range of literature reviews resulted in having bulk information about rework, as shown in Table 3.3. In the beginning, It was comprehensively investigated the causes of reworks but remain limited to explain details (Abdul-rahman, 1993; Josephson and Hammarlund, 1999). However, it was a landmark effort for other researchers to classify the groups of the leading causes. Love and Li (2000a; 2000b) started to evaluate attributes and accumulate them in a group of design, construction, project management, etc. Other researchers followed this concept, and they initiated to delve into details to capture hidden attributes behind the picture and find other attributes presented in Table 3.3.

Table 3.2. *Details of the panelists*

Title	Academic Title	Experience	Certificate
Academic Staff / Civil Engineer	Prof.	20-25	-
Mech. Eng. / Quality Cont. Man.	M.S.	20-25	(Ed2-1,2)
Mech. Eng. / Quality Cont. Man.	B.S.	15-20	(Ed2-1,2)
Mech. Eng. / Quality Cont. Man.	B.S.	15-20	(Ed2-1,2)
Architect / Quality Cont. Man.	B.S.	10-15	(Ed2-1,2)
Academic Staff / Architect	Assoc. Prof.	10-15	-
Civil Eng. / Project Manager	B.S.	25-30	-
Architect / Project Manager	B.S.	20-25	-
Architect / Quality Cont. Sup.	M.S	15-20	(Ed2-1)
Civil Eng. / Quality Cont. Sup.	Ph.D	15-20	(Ed2-1,2)
Architect/Site Eng.	B.S.	10-15	(Ed2-1)

The bulk data was consolidated before the start of the Delphi process. First, the rework cases were shared with the participants to get their attribute ideas for explaining them. Instead of scoring, they only pointed their considerations. Next, they were collected and an attribute list that spanned both literature knowledge and participants' first opinions was obtained. A second questionnaire that asked to rank the attributes prepared, and the first round of the Delphi process was initiated.

Table 3.3. *Bulk information for rework attributes*

ID	Rework attributes	Study
a1	Poor Ground Condition	
a2	Difficulty in building	Abdul-Rahman
a3	Design/Information Problems	(1993)
a4	Materials	
b1	Construction Related Problems	
b2	Design Problems	Josephson and
b3	Poor Site Management	Hammarlund
b4	Poor Workmanship	(1999)
b5	Subcontractors Problems	
c1	Change on design/construction phases	
c2	Error on design / construction phases	
c3	Omission on design/construction phases	
c7	Damage on construction	
c8	Value management	
c9	Ineffective use of IT by a design team	
c10	Design Scope freezing	Love and Li
c11	Client change	(2000a; 2000b)
c12	Poor Morale	
c13	Conflict	Love (2002)
c14	Delusion of Supervision	
c15	Contractual Claims	
c16	Cost Overruns	
c17	Time Overruns	
c18	Cost/Schedule Growth	
c19	Safety	

Table 3.3. *(Continued)*

ID	Rework attributes	Study
d1	Design Changes	Love and Edwards (2004)
d2	Construction Changes	
d3	Client	
d4	Design Team	
d5	Site Management	
d6	Subcontractor	
d7	Project Scope	
d8	Contract Documentation	
d9	Project Communication	
d10	Procurement Strategy	
d11	Design Management	
e1	Poor site condition	Yap et al. (2017)
e2	Insufficient time for the design stage	
e3	Poor coordination between client and design team	
e4	Client-related factor	
e5	Poor site supervision and inspection	
e6	Improper construction technology	
e7	Improper handling of material and delivery	
e8	Improper handling of machines and equipment	
e9	Poor contract documentation	
e10	Poor client and end-user coordination	
e11	Poor sub-contractor management	
e12	Poor site management	
e13	Construction errors due to misunderstanding of design	
e14	Poor coordination among the design team	
e15	The unclear project management process	
e16	Poor quality management by a design team	
e17	Poor quality management by the contractor	

Table 3.3. *(Continued)*

ID	Rework attributes	Study
f1	Improper handling, delivery, or providing proper materials	Balouchi et al. (2019)
f2	The unclear project management process	
f3	Poor sub-contractor management	
f4	Poor design constructability	
f5	Poor site supervision and inspection	
f6	Need for combining hard and delicate operations	
f7	Failure to define standard executive procedures	
g1	Lack of coordination and poor communication	Trach et al. (2019)
g2	The design change is initiated by the owner	
g3	Lack of experience and knowledge of the design and construction process	
g4	Lack of funding allocated for site investigations	
g5	Lack of client involvement in the project	
g6	Insufficient time and money spent on the briefing process	
g7	Expenditure on low fees for preparing contract documentation	
g8	Incomplete design at the time of tender	
g9	Poor coordination of design	
g10	The design change is initiated due to financial and economic changes	
g11	Omissions of items from the contract documentation	
g12	Errors made in the contract documentation	
g13	Insufficient time to prepare contract documentation	
g14	Inadequate client brief to prepare detailed contract documentation	
g15	Insufficient skill levels to complete the required task	
g16	Ineffective use of information technologies	

The participants were asked to rank them from one to seven. Since the sample size was limited with the panelist number, the sample mean and sample standard error were calculated instead of the population mean and standard deviation using Equations 1 and 2.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} X_i \qquad (1)$$

$$s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{x})} \qquad (2)$$

Where n represents the number of responses for the individual question, and Xi is the ranking results.

The responses were classified into three categories as indicated in the second step of Delphi. The attributes with low scores were removed from the list. Next, the second round was charged to increase the participants' consensus for remaining rework attributes. This was the case for having high standard deviations, although the mean of responses was high. A high standard deviation indicates the fluctuations in the final decision. It should be eliminated until the consensus is set. At the end of having consensus, the refined attribute for reworks was ready for further steps.

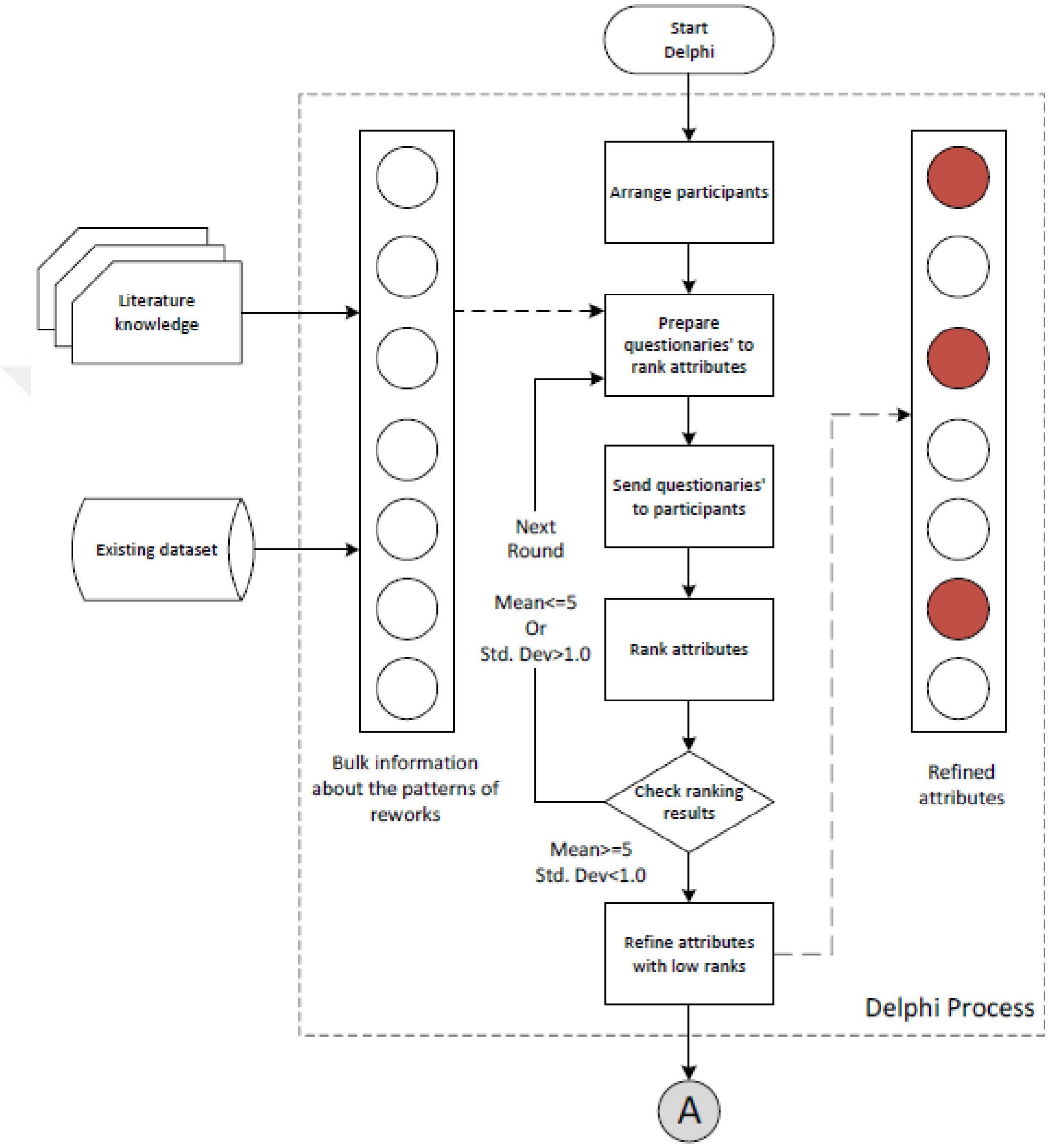

Figure 3.1. *Data preparation plan*

3.2. Determining Attribute Weights

Determining the weights of attributes is a vital step for the CBR process to be obtained accurate results. The proposed CBR models were employed either a computerized model or an expert system to determine the weights of attributes. According to Doğan et al. (Doğan et al. 2006), an automated algorithm should be implemented since selecting and finding proper experts is challenging, and these experts can be subjective. On the other hand, An et al. (An et al. 2007) were asserted that computerized systems could not understand the procedure and experts should be involved in the attribute determination process. Therefore, in this thesis, the difference between an expert system and a computerized system was investigated by employing AHP as an expert system and GA as an automated system.

3.2.1. Analytic Hierarchy Process

The weight calculations were made in two different branches. The first method preferred as an experts system is Analytical Hierarchy Process (AHP), which is the most favorable and applicable decision-making mechanism used in the literature (Alonso and Lamata 2006; Badri et al. 2012; Saaty 2008). This method's logic is based on a pairwise comparison of alternatives; therefore, the introduced strategy aims to capture the best choice. AHP is mostly conducted with the expert's opinions, but some researchers utilized the alternative observation rate in comparison (e.g., Ayhan and Tokdemir 2019b).

The decision-makers can perform straight forward ranking to select the best option, of course, but this raises a significant bias that significantly affects the final decision. AHP brings an essential advantage of reducing the inconsistency of expert opinion by proposing a solid structure for pairwise comparison (Aminbakhsh et al. 2013) and introduces control indices, as consistency index (CI) and random consistency index (CR). Saaty described the AHP steps as follow:

- Identify the problem and build up the decision hierarchy from top to goal.
- Construct the comparison matrix by following Table 3.4. C accounts for the comparison matrix in the equation, where the alternative was compared in a pairwise manner. All elements in matrix C should be higher than zero, and the dot product of elements having transverse indices results in one.

$$C = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \quad for\ \ \forall a_{ij},\ \ a_{ij} > 0;\ \ a_{ij} * a_{ji} = 1$$

$$\forall\ i, j, k = 1, 2, 3 \dots n \tag{3}$$

- Compute the sum of the comparison result for each column and call s_i. Then, the matrix C is normalized regarding the s_i values to obtain weight using Equation 4. The new matrix is called B.

$$B = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix},\ where\ \ \ b_{ij} = \frac{a_{ij}}{s_i}\ \ \ \ and\ \ \ \ s_i = \sum_{j=1}^{n} a_{ij}$$

$$\forall\ i, j = 1, 2, 3 \dots n \tag{4}$$

- The elements of B are added at each row and divided to the number of alternatives (n) to find out the weight of alternatives (see Equation 5).

$$w = \begin{bmatrix} w_i \\ \vdots \\ w_n \end{bmatrix},\ \ where\ \ \ w_i = \frac{\sum_{j=1}^{n} b_{ij}}{n}$$

$$\forall\ i, j = 1, 2, 3 \dots n \tag{5}$$

    - Two incremental criteria control the consistency in AHP. First, an average of each row in matrix B, C's normalized matrix regarding s_i, should be taken to obtain the weight vector, w. The dot product of matrix w and matrix C need to be found and normalized with the weight. The result creates a matrix R, whose maximum value, λ_{max}, designates the divergence among the attributes compared. The consistency index (CI) and consistency ratio (CR) is determined regarding Equation 6 as follows. Saaty (1990) proposed a random consistency index table (RC) given in

Table 3.5. The RC values will be determined regarding the number of alternatives that will be compared.

$$R' = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nn} \end{bmatrix} . \begin{bmatrix} w_i \\ \vdots \\ w_n \end{bmatrix} = \begin{bmatrix} r_i \\ \vdots \\ r_n \end{bmatrix}$$

$$R'_{normalized} = \begin{bmatrix} r'_i \\ \vdots \\ r'_n \end{bmatrix}, \quad where \quad r_i = \sum_{i=1}^{n} \frac{r_i}{w_i}, \quad \forall\, i = 1, 2, 3 \ldots n$$

$$\lambda_{max} = \max\,(R'_{normalized})$$

$$CI\ (Consistency\ Index) = \frac{\lambda_{max} - 1}{n-1}$$

$$CR\ (Consistency\ Ratio) = \frac{CI}{RI} \tag{6}$$

Table 3.4. *AHP Scale*

Scale	Definition	Reciprocals
1	The equal importance of two elements	1
3	Low importance of one element over another	1/3
5	Strong importance of one element over another	1/5
7	Very strong importance of one element over another	1/7
9	The absolute importance of one element over another	1/9
2,4,6,8	Intermediate values	1/2, 1/4, 1/6, 1/8

Table 3.5. *LI values proposed by Alonso-Lamata (2006)*

Element Size	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
RI	0	0	0.52	0.88	1.11	1.25	1.34	1.41	1.45	1.49	1.51	1.54	1.55	1.57	1.58

The process expressed above was integrated into the current study, as displayed in Figure 3.2. Each case was re-explained with the rework attributes, which will be given in the discussion section in detail and converted to the binary system.

If the rework-attribute is observed in cases, it will be assigned with one; otherwise, it will equal zero. The attributes were classified into four groups as Material, Operation, Construction, and Design. As mentioned before, the cases also had their cost impacts, and they were considered while performing the AHP process.

First, the main groups were compared to determine each group's contribution to reworks. Then, all items under these groups were exposed to pairwise comparison to being assigned with their weights. The group weight determination formed the first comparison that utilized the cost impacts. The high-cost impact (equal or higher than 3) controlled the comparison criteria. The high-cost impact cases were retrieved from the data, and their attributes were segmented regarding the main groups defined above. Identifying the count of occurrence followed the previous step, and it will be assigned for each group as a comparison criterion on AHP. Next, the ranking procedure started regarding this. The occurrence rates were divided into all pairwise comparison steps and resulted in insufficient numbers. These numbers will be normalized according to the expression in Table 3.4 to eliminate the vagueness.

When the first round of AHP assigned the group weights, the same process was carried out for attributes under these four groups. Unlike the previous part, each item's observation rate (f_{ij}) will be aggregated concerning the cost impact (ci) to find the weighted sum (f_{wi}) for all attributes as shown in Equation 7. Then, the weighted sum values were normalized ($f_{wi\text{-}norm}$) into the groups to make their sum equal to one. The weighted sum of each attribute was multiplied with the frequency (f_i) to set the comparison criteria (cc_i) before AHP. Details about the calculations will be given in the discussion part for a better understanding.

$$fw_{i-norm} = \frac{\sum_{i=1}^{n} fw_i}{n} \qquad ,where \quad fw_i = \sum_{j=1}^{5} c_i * f_{ij}$$

$$cc_i = f_i * fw_{i-norm} \qquad ,where \quad f_i = \sum_{j=1}^{n} f_{ij} \qquad \forall\, i,j = 1,2,3 \dots n \qquad (7)$$

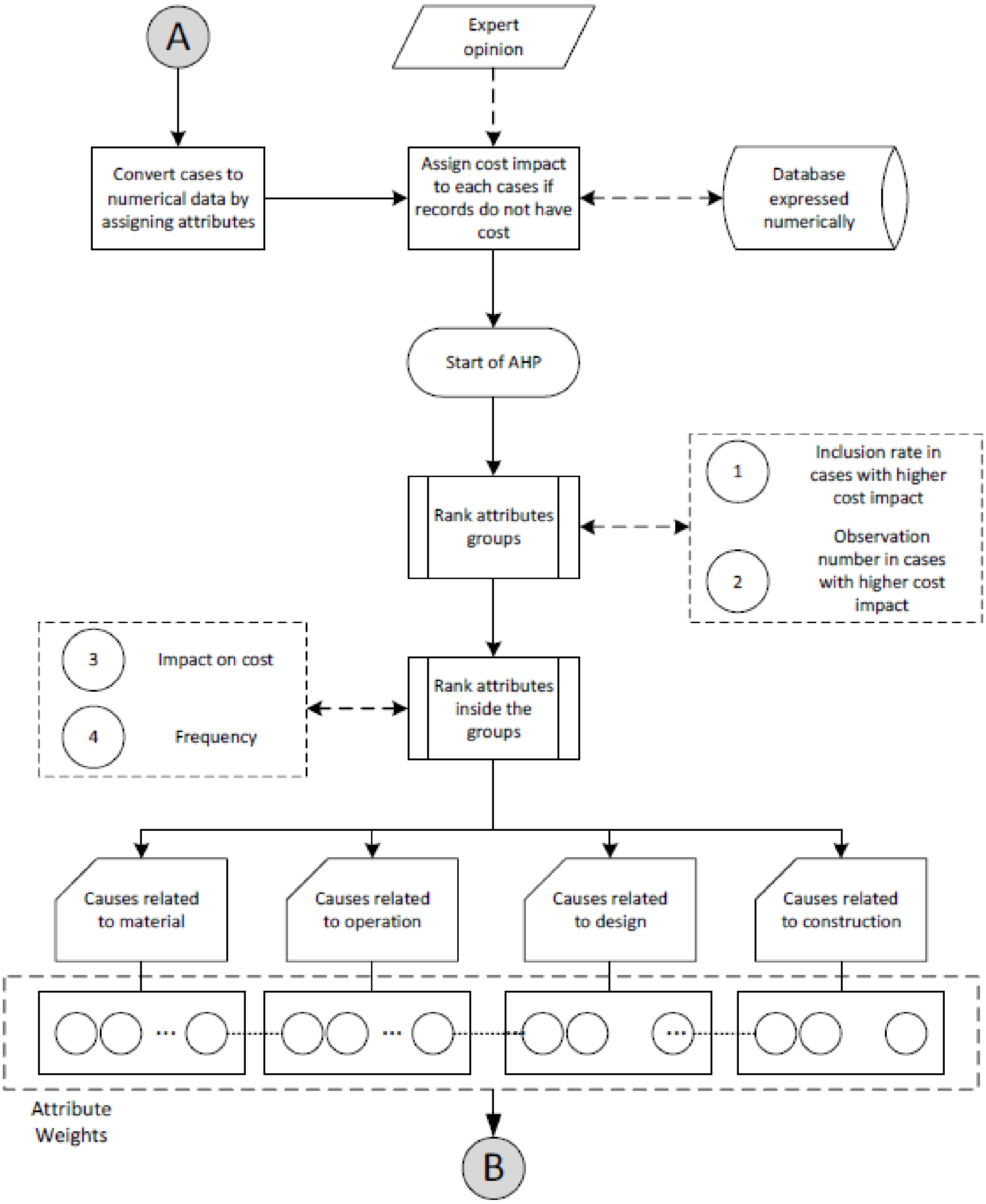

Figure 3.2. *Weight calculation by AHP*

3.2.2. Genetic Algorithm

Besides AHP, Genetic Algorithm was implemented as an automated system. GA is a heuristic AI method that focuses on obtaining near-optimum solutions for complex problems. It adopts the principles of natural selection (Doğan et al. 2006). In GA, the chromosome term is referred to as the solutions. Fitness criteria indicate the success of the chromosomes. GA process commences with a random population of chromosomes, and the success of these chromosome populations is determined. Then, in order to obtain more successful generations similar to natural selection, new chromosomes are produced with crossover and mutation, and the fitness of new chromosomes is calculated. The process is repeated until the most suitable population of the chromosome is obtained.

In this thesis, GA was employed to determine the attribute weights used in the CBR model. It was aimed to minimize the error calculated with the Root Mean Square Error (RMSE) by Equation 8. In the beginning, attribute weights were selected as 1 for all attributes, and the RMSE is calculated. Then, the GA algorithm was run, and the solution with the lowest RMSE was integrated into the CBR model. Moreover, to GA, *Evolver* from *Decision Tools Suite* (Palisade 2020) was used for this study.

$$\text{Root Mean Square Error (RMSE)} = \sqrt{\frac{\sum_{1}^{n}(y_i{}' - y_i)^2}{n}} \tag{8}$$

Where y_i' is the actual cost impact, y' is the predicted one.

3.3.CBR

Artificial intelligence techniques are widely implemented in construction industry-related researches to propose innovative solutions for the problems. Most of these techniques are rule-based; however, generally, these problems cannot be solved by sticking to a rule. Instead, better results can be obtained with knowledge-based approaches due to the complexity of construction problems.

Case-based reasoning is a knowledge-based system that is inspired by human memory and reasoning. Human memory and reasoning tend to use previous experiences when encountered with a new problem. Similarly, CBR adapts the previously proposed solutions to a new problem. Moreover, it enables to revise and improve the solution offering a mechanism by integrating the new solutions to the system. This is represented as a cycle composed of four main steps, namely retrieve, reuse, revise and retain.

As presented in Figure 3.3, CBR utilizes from the attribute weights and the cycle commences with introducing a new case to the case base. An algorithm calculates the similarity scores of cases and finds similar cases accordingly. The solutions of the most similar cases are recalled from the database. If the retrieved cases do not fulfill the requirements, the solution is revised to obtain more relevant results. Finally, the real solution of the new case is retained, and the case base is updated. Hence, CBR is an effective method to propose solutions and can revise and improve itself.

The cases are defined with the attributes. In this study, each NCR is considered as a case. The attributes were determined with Delphi Method, and the data was formed as binary variables. When a new case is introduced to the model, firstly, the attributes of the new case and the cases in the dataset were compared, and a similarity matrix was established by analogous to x-Nor operation If the status of an attribute at both the new case and the cases in the dataset is the same, then the similarity matrix's relevant element will be 1. Otherwise, it will be 0. This process was applied to all cases in the dataset.

Each attribute has a different contribution to non-conformities, and this contribution is expressed with attribute weights in CBR. Indeed, assigning attribute weights is the most crucial factor that determines how successful the CBR model is. Genetic Algorithm and Analytic Hierarchy Process were employed to specify the attribute weights. GA is a computerized method inspired by nature; on the other hand, AHP is an expert-based one. Therefore, these two methods were also used to compare the impacts of computerized and expert-based methods on the CBR model's accuracy.

Two weight matrixes were established with the results obtained from GA and AHP. Then, the similarity score of each case in the dataset was calculated with the Equation 9. In order to express the similarity score in terms of percentage, it is divided by the summation of the attribute weights. Therefore, the similarity score has a range between 0% and 100%.

$$\text{Similarity Score } = \frac{S\ x\ W}{S\_W} \tag{9}$$

where S is the similarity matrix, W is the weight matrix, and S_W is the total of attribute weights

The cases with a similarity score of at least 97% were named as similar cases. The similar cases were extracted to form a similar case matrix. In order to calculate the occurrence probability of the cases in terms of cost impact, successive cases of similar cases are needed to be obtained. However, the error and standard deviation vary depending on the selected number of successive cases. Therefore, an analysis was conducted to find the optimum number of successive cases between the range of 1 and 25. Mean Absolute Error (MAE) was calculated with the Equation 10. MAE and standard deviation were recorded and compared.

$$\text{Mean Absolute Error (MAE)} = \frac{1}{n} \sum_{t=1}^{n} | y_i' - y_i | \tag{10}$$

where y_i' is the actual cost impact, y' is the predicted one.

Finally, MAE and overall MAE were calculated to find out how successful the proposed model is. Moreover, the accuracy of CBR models with AHP and GA was also compared.

The script written in MATLAB software was used in this thesis to operate the CBR-based model regardless of the size of the dataset, and the datasets were imported from MS Excel. The script can calculate similarity scores of cases, extract similar cases, and compare actual and prediction results.

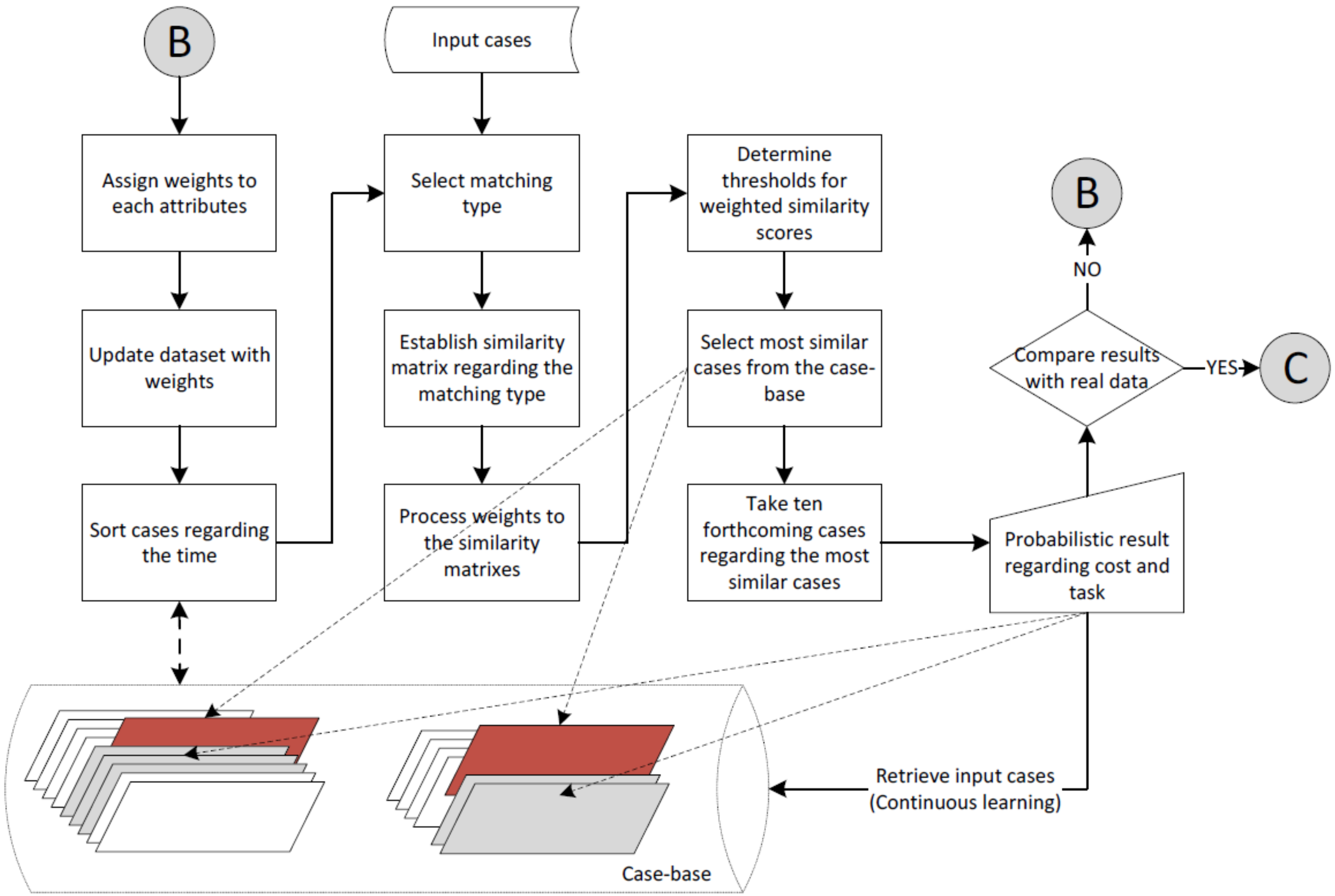

Figure 3.3. *Steps of CBR Process*

CHAPTER 4

RESULTS AND DISCUSSION

4.1.Data Preparation

The data preparation started with the literature survey, which was explained in detail. The range of rework attributes was considerably broader, so it should be refined before data modeling. The attributes taken from the literature survey were taken from Table 3.1 chronologically. As understood from the expressions, most of the attributes accounted for the same meanings. It was clear that some of them could be easily eliminated from the dataset using the author's engineering judgment. However, it was desired to obtain a coherent list of attributes that reflect both the dynamic nature of constructions and the data used for the current study. The Delphi method took place to inhere forming a list based on factual knowledge. The participants were asked to evaluate the rework cases without delivering the list in Table 3.1. However, rework cases were not entirely shared with the panelist to save time. For this reason, the cases included almost all types to ensure that panelists can interpret representative findings of all cases. After compiling the responses, the attributes in Figure 3.2 were obtained. Some of the attributes had almost the same meaning, so the authors used self-experience to refine them, as shown. For example, the term "Incorrect or defective material usage" has the same meaning as the other two terms of "Damaged material usage" and "Expired material usage". Therefore, all these three terms were decided to express as one attribute. The authors performed the same procedure for the attribute presented in Table 3.3 to confine attributes to be scored.

The early studies shown in Table 3.3 formed the bases of attribute groups. They indicated the reasons comprehensively instead of introducing a particular problem. The author reorganized these attributes into four main categories before scoring them. These attributes were associated with the types of activity and the consequences when rework occurred. These groups can be divided as follow: 1) Materials, 2) Operational, 3) Design, and 4) Construction. The nonconformance data constrained the scope of the thesis by considering only construction failure. The records were taken during the construction phase, so they did not compose any loss occurring in the design or tendering phases. Therefore, client-related and subcontractor-related factors mainly were eliminated. Some of them participated in the Delphi process to be scored whether it has an impact on the cases or not. Besides, it was expected that the rework attributes, including the pre-construction process, were removed from the attribute list after applying Delphi since the nonconformance list focused on the reworks during the construction phase only.

As mentioned before, the responses coming from panelists combined with them the literature knowledge were given in Table 3.3 and the refined list was shown in Table 4.1, which was sent to panelists for scoring them. Then, the first round was kicked off.

Table 4.1 underlined each response of the panelist in detail. They were asked to rank the attributes from one to seven. The one accounted for the strongly disagree with adding the relevant attribute into the final list, whereas the seven stands for expressing the strongly agree term. At most, two rounds were performed since the panelists came to common ground at the end of the second round. In other words, the panelists achieved the consensus among them. When the first round was completed, the results were interpreted. As indicated in the methodology section, the results were interpreted into three categories. If the ranking score ends with a high score with a significant standard deviation, the attribute will remain for a second round. Some attributes remarked with "*" need to be discussed in further rounds since the significant fluctuation existed among the panelist's ranking scores.

For example, "Damaging material during transportation/loading" required a second round because it got a considerably high standard deviation. Panelists 1 and 2 ranked a low score, but they slightly directed their thoughts over consensus ground in the second round, and it was accepted as a rework attribute. The same thing occurred for "Poor quality management by design team" and "Design freezing scope." However, panelists were inclined to disagree with these attributes, so they were not accepted to the rework attribute list accordingly. In general, design-related attributes were not accepted since the rework cases did not cover design-related nonconformances, as indicated at the beginning.

Moreover, panelists were allowed to indicate their thoughts on the questionnaire. Unlike the other panelists, Panelist 1 and 9 put a tremendous recommendation to express some attributes. They claimed that "Inadequate Training, Inadequate Staff and Insufficient / Unproper Workmanship" can be expressed as one term under the latest attribute. They implied that the meaning of the third attribute encapsulated the first two. This comment was evaluated and decided to use a single term of "Inadequate Staff and Insufficient / Unproper Workmanship," as panelists defined. At the end of the second round, the remaining also agreed on this and accepted. However, the other comments indicated in Table 4.1 were not accepted. Some panelists suggested combining "Problem with warehouse (Labeling etc.) with the "Problems with documentation." This comment was not considered well because they considered that the documentation problem extends beyond the labeling or other issues encountered in the warehouse. Ultimately, the list of attribute s shaped its last version, as indicated in Table 4.2.

After the attributes were determined, the data was modelled by using them. For better understanding, the case with ID of 88 was examined. The explanation of the NCR is that the wall was seriously damaged during the installation of instruments and the cost impact of this case was assigned as 5. This non-conformance includes the attributes with the IDs of C1, C6 and O4. In detail, since the walls were completed product and they were damaged during another installation operation, C1 attribute which is the “Damaging the Completed Work” was included. Moreover, this non-conformance can be avoided with proper supervision and by implementing the specifications properly; therefore, C6 and O4 attributes which are “Incompliance with Technical Specification” and “Lack of Supervision” were also involved. While converting the data into binary format as above-mentioned example case, the involving attributes were determined, and the corresponding value of dataset was set as 1 and otherwise it will be 0. This procedure was repeated for all the cases in the dataset.

Table 4.1. *Delphi results*

Group	Attributes	Round 1													Notes	Round 2												
		Panelist											μ	σ		Panelist											μ	σ
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11				P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11		
Material related	Improper handling of machines and equipment	4	4	3	3	3	5	7	4	2	2	2	3,5	1,5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Improper handling of material and delivery	5	6	6	6	5	6	6	7	6	6	6	5,9	0,5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Incorrect or Defective Material Usage	5	5	6	6	6	7	7	5	6	6	7	6	0,8	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Procurement of Incorrect Material	5	5	5	7	7	7	6	7	6	5	5	5,9	0,9	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Damaging Material During Transportation/Loading*	4	3	5	5	5	5	6	6	7	6	7	5,4	1,2	Second round is required	5	6	5	5	5	6	5	6	7	6	7	5,7	0,8
Design	Design problem / changes on construction	6	6	6	6	6	6	6	7	6	6	6	6,1	0,3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Omission on design	2	1	2	2	2	3	4	1	1	1	1	1,8	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Design Team / Management	1	1	1	1	1	3	4	1	1	1	1	1,5	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Poor coordination between client and design team	2	1	1	1	1	1	3	2	1	1	1	1,4	0,7	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Poor coordination among design team	1	1	1	1	1	1	1	2	1	2	2	1,3	0,5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Poor quality management by design team*	3	4	5	4	4	5	4	4	4	5	5	4,3	0,6	Second round is required	3	4	5	4	4	5	4	4	4	5	5	4,3	0,6
	Ineffective use of IT by design team	1	1	1	1	1	1	1	1	1	1	1	1	0	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Design Scope freezing*	4	4	3	4	3	3	3	2	3	2	3	3,1	0,7	Second round is required	4	4	3	4	4	3	3	2	3	2	3	3,2	0,8
	Insufficient time for design stage	2	3	3	3	2	3	3	3	3	3	3	2,8	0,4														
Construction	Damaging the Completed Work	5	6	5	6	7	7	7	7	7	6	7	6,4	0,8	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Work in Confined Space	7	6	7	7	7	7	6	7	7	6	7	6,7	0,5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Construction errors due to misunderstanding of design	7	7	7	7	7	7	7	5	5	5	5	6,3	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Inadequate Preparation before Starting the Work*	5	5	5	4	5	5	6	7	7	4	5	5,3	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Inadequate Site Cleaning after Completing the Work*	5	6	6	5	4	4	4	5	5	5	6	5	0,8	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Incompliance with Technical Specification	6	7	6	6	6	6	6	6	5	7	7	6,2	0,6	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Insufficient Training*	6	6	5	5	6	6	6	6	6	6	6	5,8	0,4	Panelist 1 and 9 recommended to combine these 3 attribute "Insufficient / Unproper Workmanship" - "Yes"	-	-	-	-	-	-	-	-	-	-	-	-	-
	Inadequate Staff*	1	5	5	5	6	6	6	6	1	6	6	4,8	1,9		-	-	-	-	-	-	-	-	-	-	-	-	-
	Insufficient/Unproper Workmanship*	1	5	6	5	6	6	7	7	1	6	7	5,2	2,2		7	7	7	7	7	6	7	7	7	6	7	6,8	0,4

Table 4.1. *(Continued)*

Group	Attributes	Round 1													Notes	Round 2												
		Panelist											μ	σ		Panelist											μ	σ
		P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11				P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11		
	Lack of Documents on Site	5	5	5	6	6	5	6	6	5	5	5	5,4	0,5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Inadequate Tools/Equipment	6	6	6	5	6	6	7	7	7	7	7	6,4	0,7	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Inadequate Application Procedure	7	6	6	6	7	7	7	7	7	7	7	6,7	0,5	Panelist 1 recommended to combine this with the previous attribute - "No"	-	-	-	-	-	-	-	-	-	-	-	-	-
	Insufficient Review of Drawings*	5	6	6	6	7	7	7	4	4	5	4	5,5	1,2		5	6	6	6	7	7	7	5	6	5	4	5,8	1
Construction	Lack of Drawings on Site	5	6	6	6	5	6	7	7	6	4	5	5,7	0,9	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Delays in Construction Timeline	6	6	6	6	6	6	7	6	5	6	6	6	0,4	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Insufficient number of Site Supervisor	6	6	6	6	5	6	6	6	7	6	6	6	0,4	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Not Following the Work Sequence	5	5	6	5	6	6	6	5	7	6	6	5,7	0,6	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Problems with Purchasing Department	5	6	6	6	6	7	7	7	7	7	7	6,5	0,7	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Problems with Warehouse (Labelling etc.)	5	6	7	7	7	7	7	7	7	7	7	6,7	0,6	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Sending Wrong Material from Warehouse	4	5	6	7	7	7	7	7	7	5	5	6,1	1,1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Lack of Supervision*	3	5	6	6	5	5	6	5	6	6	5	5,3	0,9		3	5	6	6	5	5	6	5	7	6	5	5,4	1
Operational	Problems with Documentation*	1	4	5	5	5	6	6	7	6	5	6	5,1	1,6	2nd attribute covers this - "No"	3	5	5	5	5	6	6	7	7	5	6	5,5	1,1
	Contract Documentation	1	3	3	2	2	2	2	2	2	2	2	2,1	0,5	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Project Communication	5	2	3	2	2	2	2	1	2	2	2	2,3	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Poor client and end-user coordination	1	1	1	1	1	1	1	1	2	1	1	1,1	0,3	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Unclear project management process	4	3	1	1	4	1	1	5	1	1	1	2,1	1,6	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	Poor sub-contractor management	1	1	1	1	1	1	1	1	1	1	1	1	0	Cases did not cover related knowledge	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 4.2. *List of rework attributes used in the study*

Group	Attributes	ID
Material Related	Improper handling of material and delivery	M1
	Incorrect or Defective Material Usage	M2
	Procurement of Incorrect Material	M3
	Damaging Material During Transportation/Loading*	M4
Design	Design problem/changes on construction	D1
Construction	Damaging the Completed Work	C1
	Work in Confined Space	C2
	Construction errors due to misunderstanding of design	C3
	Inadequate Preparation before Starting the Work*	C4
	Inadequate Site Cleaning after Completing the Work*	C5
	Incompliance with Technical Specification	C6
	Insufficient/Unproper Workmanship*	C7
	Lack of Documents on Site	C8
	Inadequate Tools/Equipment	C9
	Inadequate Application Procedure	C10
	Insufficient Review of Drawings*	C11
	Lack of Drawings on Site	C12
	Delays in Construction Timeline	C13
	Insufficient number of Site Supervisor	C14
	Not Following the Work Sequence	C15
Operational	Problems with Purchasing Department	O1
	Problems with Warehouse (Labelling etc.)	O2
	Sending Wrong Material from Warehouse	O3
	Lack of Supervision*	O4
	Problems with Documentation*	O5

4.2. Calculation of Attribute Weights

4.2.1. AHP

After assigning the rework attributes found in the previous chapter as presented in Table 4.2, the study advanced with the weight calculation. As indicated before, two criteria regulated the pairwise comparison. These were the frequency of the attributes in the database (f_i) and the weighted frequency of attributes (f_{wi}). Table 4.3 corresponds to the related information through the process, but it excludes the corresponding design attribute. The reason is that the design group included only a single attribute, so the author did not intend to share to avoid abundance in Table. Thus, Table 4.3 classified the attributes with respect to the observation rate for each cost impact (ci) separately. Then, the steps introduced in Equation 7 were followed to determine the normalized weighted frequency used for pairwise comparison with each attribute's frequency (f_i).

The reason for considering cost is to prevent eliminating the impact of crucial attributes that significantly influence the cost. As indicated in Table 4.3, although some attributes were highly observed through the cases, they did not cause high-cost rework (e.g., M-2, "Incorrect or Defective Material Usage"; O-4, "Lack of Supervision"). However, some of them resulted in severe damage even though the observation rate of them was rare. While pairwise comparison, considering only frequency might mislead the final decision, so the author did not pass over its effect.

As mentioned before, AHP applied two different levels, and Figure 3.2 exhibits these steps. The literature knowledge and the dataset led to classifying the rework attributes into four major groups. The first layer AHP involved the group comparison, so each rework attribute group was assigned weight. At the end of the first layer AHP, the construction-related attributes took the lead, and attributes under operation followed. Since the NCRs recorded were taken from the construction sites, the impact of design-related attributes was not significant as expected.

Table 4.3. *Details of rework attributes before AHP*

Attribute groups	Cost Impact (ci)	Attributes ID number 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Causes related to inventory Group Name (M)	1	0,516	0,571	0,565	0,395											
	2	0,320	0,298	0,391	0,395											
	3	0,118	0,090	0,022	0,158											
	4	0,039	0,036	0,022	0,053											
	5	0,007	0,006	-	-											
fw_i		1,699	1,607	1,500	1,868											
$fw_{i\text{-norm}}$		0,255	0,241	0,225	0,280											
f_i		0,198	0,693	0,060	0,049											
Causes related to procurement Group Name (O)	1	0,714	0,250	-	0,603	0,711										
	2	0,214	0,500	-	0,268	0,111										
	3	0,071	0,167	1,000	0,087	0,141										
	4	-	0,083	-	0,036	0,037										
	5	-	-	-	0,005	-										
fw_i		1,357	2,083	3,000	1,571	1,504										
$fw_{i\text{-norm}}$		0,143	0,219	0,315	0,165	0,158										
f_i		0,010	0,009	0,001	0,880	0,100										
Causes related to construction Group Name (C)	1	0,529	-	0,637	0,649	0,623	0,590	0,617	0,767	0,500	0,670	0,582	0,615	0,250	0,182	0,563
	2	0,303	0,667	0,254	0,286	0,264	0,268	0,277	0,167	0,225	0,247	0,286	0,269	0,450	0,545	0,291
	3	0,108	-	0,066	0,039	0,057	0,088	0,072	0,067	0,225	0,049	0,084	0,077	0,200	0,273	0,107
	4	0,044	0,333	0,036	0,013	0,057	0,044	0,028	-	0,025	0,026	0,041	0,038	0,100	-	0,029
	5	0,017	-	0,008	0,013	-	0,010	0,007	-	0,025	0,008	0,007	-	-	-	0,010
fw_i		1,717	2,667	1,524	1,455	1,547	1,615	1,531	1,300	1,850	1,455	1,605	1,538	2,150	2,091	1,631
$fw_{i\text{-norm}}$		0,067	0,104	0,059	0,057	0,060	0,063	0,060	0,051	0,072	0,057	0,062	0,060	0,084	0,081	0,064
f_i		0,056	0,001	0,321	0,014	0,010	0,077	0,200	0,011	0,007	0,192	0,082	0,005	0,004	0,002	0,019

over	Material	Operation	Design	Construction
Material	1	0.5	5	0.167
Operation	2	1	8	0.333
Design	0.2	0.125	1	0.111
Construction	6	3	9	1
W	0.134	0.247	0.0403	0.5782

over	M1	M2	M3	M4
M1	1	0.333	4	4
M2	3	1	9	9
M3	0.333	0.111	1	2
M4	0.250	0.111	0.50	1
W	0.24	0.62	0.085	0.056

Material

over	O1	O2	O3	O4	O5
O1	1	0.50	6	0.111	0.111
O2	2	1	8	0.111	0.125
O3	0.167	0.125	1	0.111	0.111
O4	9	9	9	1	9
O5	9	8	9	0.111	1
W	0.069	0.096	0.027	0.549	0.26

Operation

Reciprocals

over	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15
C1	1														
C2	0.25	1													
C3	2	6	1												
C4	0.5	2	0.5	1											
C5	0.5	2	0.5	0.5	1										
C6	2	4	0.5	2	2	1									
C7	2	5	0.5	2	2	2	1								
C8	0.5	2	0.5	0.5	0.5	0.5	0.5	1							
C9	0.5	2	0.5	0.5	0.5	0.5	0.5	0.5	1						
C10	3	5	0.5	2	2	2	0.5	2	2	1					
C11	2	4	0.5	2	2	2	0.5	2	2	0.5	1				
C12	0.5	2	0.25	0.333	0.5	0.5	0.333	0.5	0.5	0.5	0.5	1			
C13	0.5	2	0.25	0.333	0.5	0.5	0.5	0.5	0.5	0.5	0.5	2	1		
C14	0.5	3	0.2	0.5	0.5	0.5	0.25	0.5	0.5	0.25	0.5	0.5	0.5	1	
C15	0.5	2	0.50	2	2	0.5	0.5	2	2	0.5	0.5	2	2	2	1
W	0.073	0.021	0.137	0.064	0.054	0.080	0.115	0.050	0.045	0.106	0.088	0.033	0.38	0.030	0.065

Construction

Figure 4.1. *Pairwise comparison of rework attributes*

The second-layer AHP process elaborated the weight assignment process. This process was applied for each primary group of rework except for the design group. The reason is that there was only one attribute under the design group, so the weight of the design group was directly equal to the attribute's weight.

Figure 3.2 demonstrates the second-layer AHP process by extracting additional boxes from the main comparison matrix. While conducting a comparison, f_i and f_{wi} were considered together and assigned to all rows and columns. Then, the comparison was initialized.

According to the results of the AHP process, each group has a dominant rework attributes. For example, "M-2, Incorrect or Defective Material Usage" was the most significant contributor for NCRs in material groups. "O-4, Lack of supervision" was the most influential attribute for operation groups. However, the weight of the attributes aggregated more uniformly under the construction groups, so it is not proper to infer anything about the most influential ones.

Eventually, Table 4.4 tabulated the AHP process results, which would be one of the Case-Based Reasoning process inputs. Consistency ratio, CR, controlled the AHP calculations to check whether the scoring alternatives are logical or not. As indicated in the methodology section, CR should be less than 10%, so each comparison step for all layers satisfied this condition.

Table 4.4. *Attribute weights after AHP*

Attribute groups	Group ID	Group Weights	CR for groups	Attributes ID number															CR'
				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Causes related to iventory	M	0,134		0,240	0,620	0,085	0,056	-	-	-	-	-	-	-	-	-	-	-	0,078
Causes related to procurement	O	0,247	0,091	0,069	0,096	0,027	0,549	0,260	-	-	-	-	-	-	-	-	-	-	0,018
Causes related to design	D	0,040		0,040	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Causes related to construction	C	0,578		0,073	0,021	0,137	0,064	0,054	0,080	0,115	0,050	0,045	0,106	0,088	0,033	0,038	0,030	0,065	0,050

4.2.2. GA

Besides AHP, GA is the selected method as an automated system for the calculation of attribute weights. *Evolver* from *Decision Tools Suite* is used for these calculations (Palisade 2020). Before commencing the GA analysis, the initial attribute weights and an equation are needed to be determined. Therefore, the process was initiated by selecting 1 for all attributes weights, and it was aimed to obtain the minimum RMSE value.

GA is an iterative process. In each iteration, the algorithm either crosses over or mutates the attribute weights. The algorithm has generated 35.872 iterations; however, it has reached the local minimum at the 26.375th iteration. Consequently, RMSE values obtained during the analysis were presented in Figure 4.2, and the calculated attribute weights as a result of GA analysis were provided in Table 4.5.

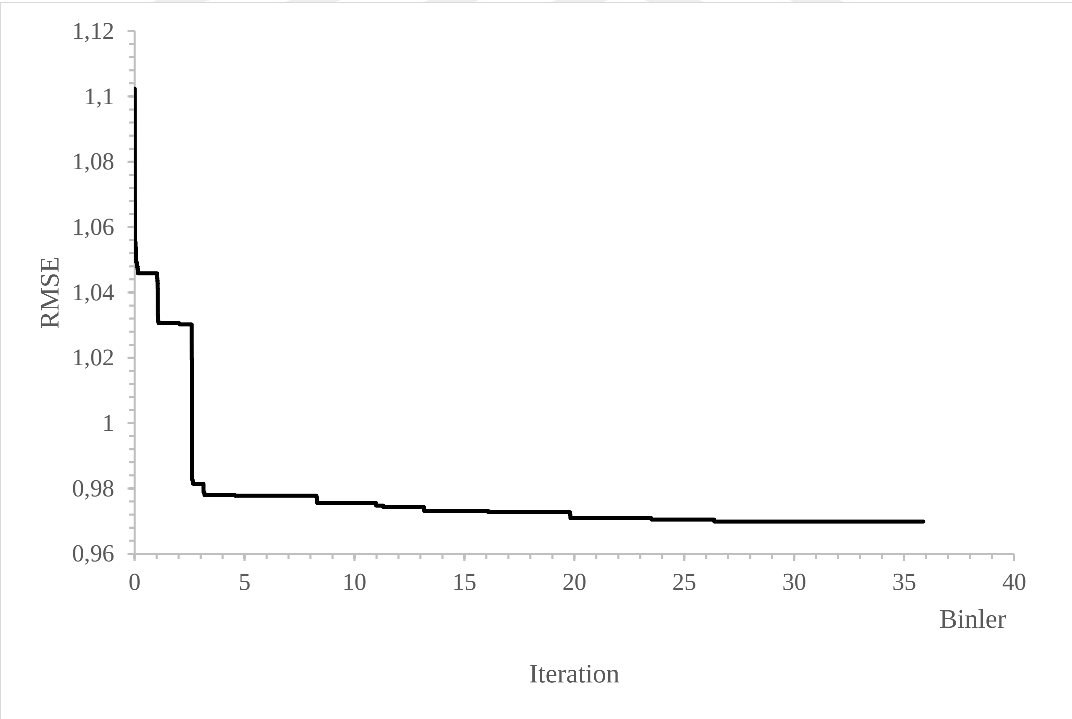

Figure 4.2. *RMSE vs. Iteration number*

Table 4.5. *Attribute weights after GA*

Attribute groups	Group ID	Attributes ID number														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Causes related to iventory	M	0,958	0,013	0,716	0,883	-	-	-	-	-	-	-	-	-	-	-
Causes related to procurement	O	0,226	0,966	0,762	0,551	0,894	-	-	-	-	-	-	-	-	-	-
Causes related to design	D	0,852	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Causes related to construction	C	0,592	0,781	0,975	0,301	0,17	0,001	0,072	0,572	0,601	0,019	0,542	0,612	0,88	0,988	0,519

4.3.CBR

CBR is an effective machine learning tool that benefits from solutions to previously experienced problems. The performance of prediction via CBR depends on different variables. Attribute weights and matching strategy were two of them. As mentioned above, AHP and GA were employed to find the attribute weights, and the exact matching strategy was decided to be implemented since the attributes are linguistic variables expressed in binary format.

In order to initiate the CBR process, 150 test cases were selected among the data. The similarity matrix for each test case was established regarding the matching status, analogous to the logical x-NOR operation, and the similarity score of each stored case was calculated. The cases with at least 97% similarity score were extracted.

The aim of the study is to estimate the cost impacts of upcoming NCRs in order to avoid negative cost impacts of NCRs on projects by informing the practitioners about the possible deficiencies in a quality system. In order to estimate the cost impacts of upcoming cases, successive cases following similar cases were also obtained. However, the model's accuracy varies with the number of how many successive cases should be involved in the analysis. Therefore, the optimum number of successive cases were analyzed within the range of 1 and 25.

As shown in Table 4.6 and Figure 4.3, the model using GA has a slightly better MAE when compared with the model using AHP. Moreover, although the minimum MAE is calculated as 7,81% and 7,22% for AHP and GA, respectively, when 3 successive cases were involved, the minimum standard deviation was obtained when the number of successive cases was 10. Since the uncertainty decreases while the deviation gets smaller, and the lowest standard deviation was obtained both for the models employing AHP and GA when the number of successive cases is 10, the successive number of cases was selected as 10.

Table 4.6. *Overall MAE and standard deviations of successive cases*

Number of Successive Cases	AHP		GA	
	Overall MAE	Standard Deviation	Overall MAE	Standard Deviation
1	13,54%	7,69%	12,09%	7,77%
2	9,83%	4,94%	8,95%	4,97%
3	**7,76%**	**4,99%**	**7,20%**	**5,06%**
4	8,39%	5,18%	7,94%	5,27%
5	7,91%	4,94%	7,53%	5,08%
6	7,93%	4,62%	7,55%	4,80%
7	8,03%	4,72%	7,54%	4,91%
8	7,79%	4,19%	7,31%	4,36%
9	8,08%	3,97%	7,53%	4,12%
10	**8,12%**	**3,93%**	**7,55%**	**4,09%**
11	8,34%	4,05%	7,80%	4,22%
12	8,44%	4,09%	7,88%	4,27%
13	8,59%	4,26%	8,03%	4,43%
14	8,60%	4,26%	8,07%	4,45%
15	8,68%	4,25%	8,11%	4,40%
16	8,78%	4,33%	8,19%	4,49%
17	8,83%	4,44%	8,25%	4,62%
18	8,96%	4,66%	8,38%	4,84%
19	9,18%	4,73%	8,60%	4,90%
20	9,26%	4,77%	8,70%	4,95%
21	9,31%	4,85%	8,75%	5,03%
22	9,26%	4,93%	8,71%	5,12%
23	9,21%	4,97%	8,67%	5,16%
24	9,26%	5,03%	8,72%	5,22%
25	9,24%	5,01%	8,69%	5,20%

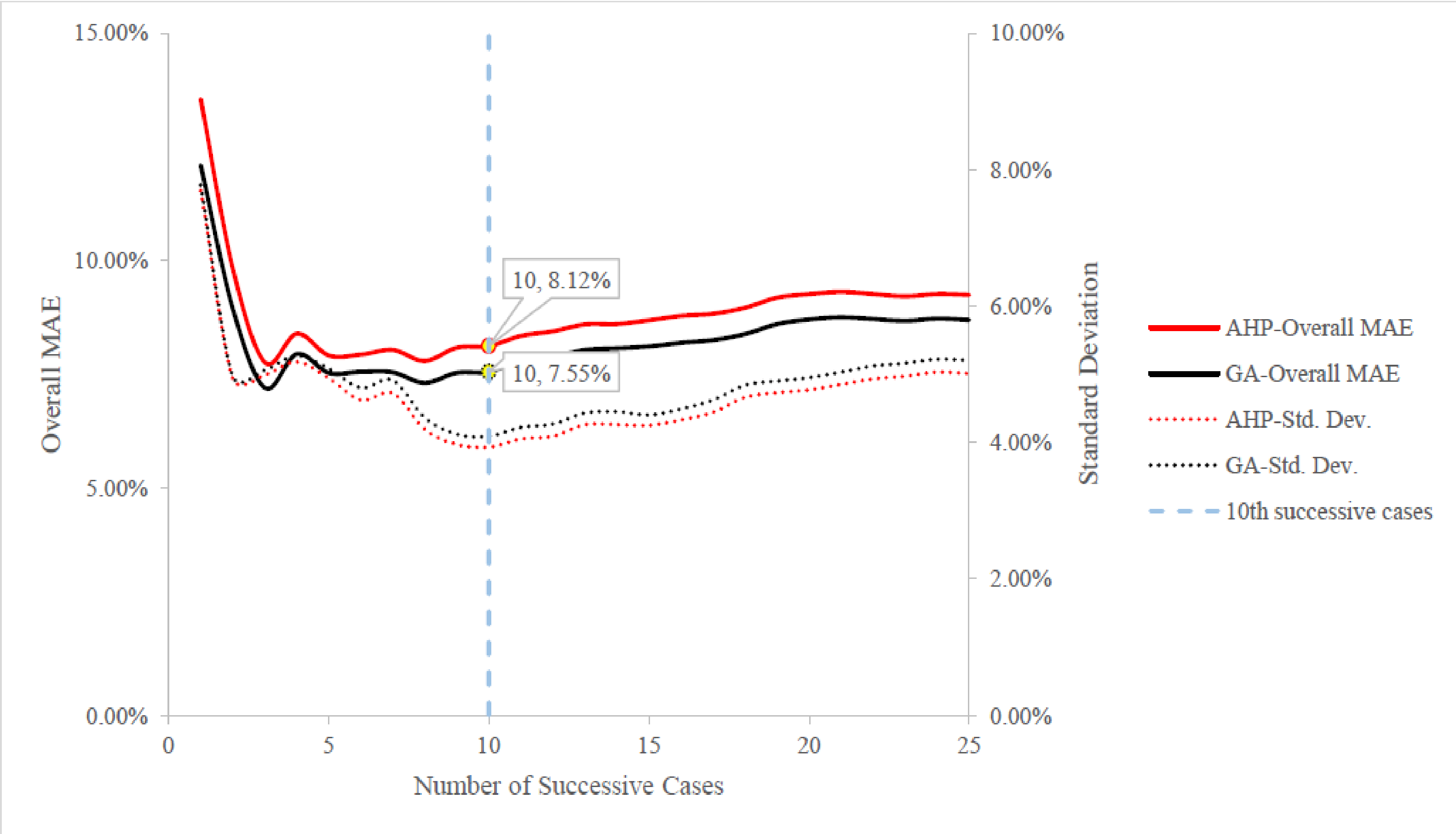

Figure 4.3. *Overall MAE vs. number of successive cases for the models employing AHP and GA*

As mentioned, the attribute weights are indicator of the impact of attributes on the non-conformances. The 3 attributes with highest weights were M2, O4 and O5 for AHP and C14, C3 and O2 for GA, respectively. Hence, according to the results of expert and automated systems, different attributes have higher significance and impact on NCRs.

The actual occurrence percentage of successive NCRs after the randomly selected cases were presented for all five cost impacts in Appendix A. For instance, After the 18th NCR recorded in the project, 2 non-conformances with cost impact of 1 were observed, and 5, 2, 1 and 0 NCRs were recorded for cost impacts of 2, 3, 4 and 5 respectively. Besides, the results after CBR process were presented in Appendices B and C for the processes employed AHP and GA, respectively. These results were the predicted ones, and in order to measure the accuracy of the developed model an error for all the randomly selected 150 cases were calculated with the Equation 10.

Another objective of the study is to determine whether an expert system or an automated system is more suitable for CBR analysis. Therefore, the results obtained from the models employing AHP and GA were compared. As shown in Table 4.6, although the average MAE of each of five cost impacts in the model using attribute weights obtained via GA is less than ones obtained via AHP, standard deviations of each five cost impacts in the AHP model are less than GA model. Secondly, while the overall MAE and standard deviation of AHP are found as 8,12% and 3,93%, respectively, they are 7,55% and 4,09 % in the GA model. Therefore, it was concluded similarly to the comparison of MAE and standard deviation values of the two models.

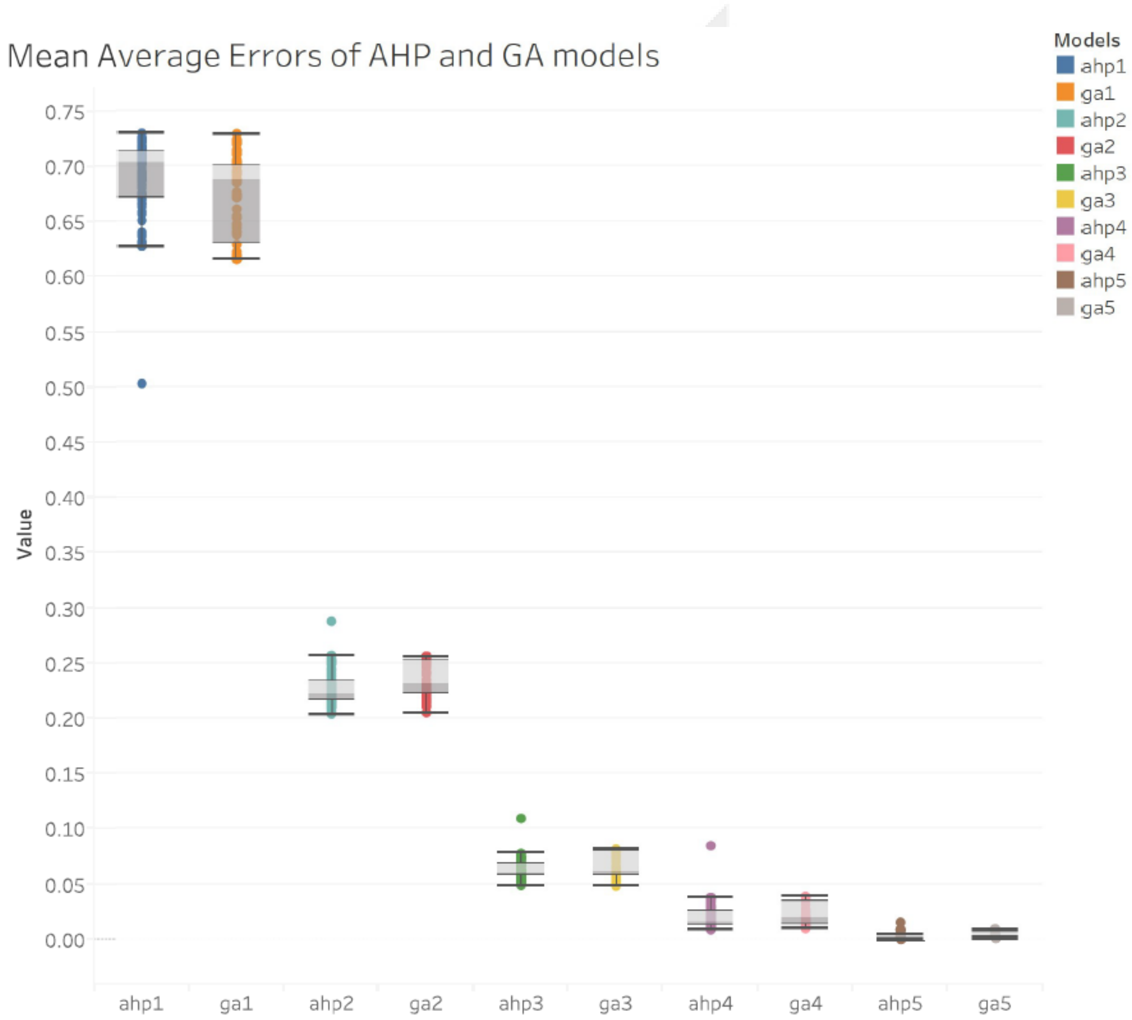

Figure 4.4. *Comparison of MAE values of AHP and GA models*

CHAPTER 5

CONCLUSION

Quality problems in construction projects may lead to negative consequences such as delays in the project timeline, cost overruns, and damage the companies' reputation. Hence, the main objective of this study is to prevent quality problems via implementing a novel predictive early warning mechanism for these problems in construction projects. Furthermore, record-keeping is a challenging and underestimated issue for construction projects. Practitioners either do not give due importance to record the project information wholly and correctly or have insufficient experience in developing and implementing efficient record-keeping systems. This situation hinders benefiting from the implementation of lessons learned from the information obtained via these records. Therefore, the study also emphasizes the significance of record-keeping and aid in eliminating the inconsistencies in record-keeping of quality problems via proposing a guide for practitioners.

The study is composed of three main steps. Firstly, quality data which was sorted according to the occurrence time was collected from a construction project, and Delphi Method was employed to determine the attributes expressing the quality problems concisely. Accordingly, a literature review was conducted to specify the leading attributes included in previous researches. It was requested from the experts indicated in Table 3.2 to rank the attributes, and it was ensured that none of the rankings of participating experts are affected by each others'. Attributes were iteratively ranked till a consensus on the final attribute list was reached, and the final list of attributes was shown in Table 4.2. Moreover, this list of attributes can be employed by quality control practitioners in a construction project in order to record the NCRs and to be involved in further and detailed analysis.

The selection of the attributes enables to express the NCRs numerically; therefore, the data was converted in binary format. However, these attributes have different contribution on the occurrence of the NCRs, and the rate of contribution was expressed with attribute weights. In order to determine the attribute weights, two different methods, AHP and GA, were implemented. Using these methods also allows observing the performances of using experts' opinions and utilizing automated systems to compare the accuracy of the predictive models. In AHP, expert opinions are involved to determine the hierarchy of the attributes, in order words, the attributes are determined by experts. However, in this study, it was utilized from the observation rates of attributes and cost impacts in AHP. The obtained results were given in Table 4.4. On the other hand, GA which imitates the process of natural selection was employed as an automated system. Attributes weights were initially selected as 1, and with the help of *Evolver* from *Decision Tools Suite* (Palisade 2020), weights which are presented on Table 4.5 were acquired.

Finally, CBR process was initiated in order to estimate the possible further outcomes of NCRs in terms of cost impacts and to warn the practitioners about them if any measure is not implemented to the quality system. CBR benefits from the principle of a similar solution can be applied to similar solutions; therefore, at the beginning of the process, 150 random cases were selected. The similarity matrix was established, and the similarity score was calculated for all these random cases by using the attribute weights obtained via AHP and GA. The cases with a similarity score of at least 0,97 were determined. In order to calculate the estimated occurrence percentage in terms of cost impact, it was necessary to investigate how many successive cases should be included in CBR. Therefore, an analysis to determine the number of successive cases was conducted within the range of 1 and 25. MAE and standard deviation of analysis of successive case numbers within the indicated range were calculated for both using AHP and GA as attribute weights, and the results were presented in Table 4.6 and Figure 4.3. It was found out that although the analysis using GA have less MAE than the analysis using AHP, the situation was vice versa when standard deviations were compared.

The study makes three main contribution to the literature. Record-keeping is a challenging and underestimated issue for construction projects. Practitioners often do not give due importance to recording project information accurately and in full or they have insufficient experience in developing and implementing efficient record-keeping systems. Thus, the study offers an insight to further research with a concise list of attributes obtained by the expert opinions. Furthermore, it was revealed that cost impacts of quality problems can be predicted by CBR, and the influence of attributes on NCRs was evaluated by implementing AHP and GA. Finally, the effectiveness of expert and automated systems was compared, and it was concluded that automated systems are as adoptable and effective as expert systems when the quality problems are concerned.

The study also has limitations as well as the provided benefits. Firstly, the data preparation method applied in this study is considerably long-lasting operation; however, it can be shortened by employing other AI methods such as Natural Language Processing. It can ease the evaluation of NCRs and data modelling. Secondly, in order to achieve accurate and applicable results, it is crucial that all the quality problems are appropriately recorded. Therefore, at the beginning of the construction projects, it can be beneficial to be trained the professionals who will run the proposed CBR model. Moreover, the relationships between the attributes should also be examined. Although attributes weights calculated as in this study gives an insight about the impact of attributes, the relationship and correlations between the attributes should also be investigates. As a future study, these relations can be examined by using ARM. Furthermore, other methods rather than AHP and GA will be employed in the calculation of attribute weights process in future studies, since better results can be obtained via using other methods.

REFERENCES

Aamodt, A., and Plaza, E. (1994). "Case-based reasoning: foundational ıssues, methodological variations, and system approaches." *AI Communications*, 7(1), 39–59.

Abdul-Rahman, H. (1993). "Capturing the cost of quality failures in civil engineering." *International Journal of Quality & Reliability Management*, 10(3), 20–32.

Ahn, J., Ji, S. H., Ahn, S. J., Park, M., Lee, H. S., Kwon, N., Lee, E. B., and Kim, Y. (2020). "Performance evaluation of normalization-based CBR models for improving construction cost estimation." *Automation in Construction*, Elsevier, 119(November), 103329.

Alonso, J. A., and Lamata, M. T. (2006). "Consistency in the analytic hierarchy process: A new approach." *International Journal of Uncertainty, Fuzziness and Knowlege-Based Systems*, 14(4), 445–459.

Aminbakhsh, S., Gunduz, M., and Sonmez, R. (2013). "Safety risk assessment using analytic hierarchy process (AHP) during planning and budgeting of construction projects." *Journal of Safety Research*, National Safety Council and Elsevier Ltd, 46, 99–105.

An, S. H., Kim, G. H., and Kang, K. I. (2007). "A case-based reasoning cost estimating model using experience by analytic hierarchy process." *Building and Environment*, 42, 2573–2579.

Arditi, D., and Tokdemir, O. B. (1999a). "Comparison of case-based reasoning and artificial neural networks." *Journal of Computing in Civil Engineering*, 13(3), 162–169.

Arditi, D., and Tokdemir, O. B. (1999b). "Using case-based reasoning to predict the outcome of construction litigation." *Computer-Aided Civil and Infrastructure Engineering*, 14(6), 385–393.

Ayhan, B. U., Doğan, N. B., and Tokdemir, O. B. (2020). "An association rule mining model for the assessment of the correlations between the attributes of severe accidents." *Journal of Civil Engineering and Management*, 26(4), 315–330.

Ayhan, B. U., and Tokdemir, O. B. (2019a). "Safety assessment in megaprojects using artificial intelligence." *Safety Science*, Elsevier, 118, 273–287.

Ayhan, B. U., and Tokdemir, O. B. (2019b). "Predicting the outcome of construction incidents." *Safety Science*, Elsevier, 113, 91–104.

Ayhan, B. U., and Tokdemir, O. B. (2020). "Accident analysis for construction safety using latent class clustering and artificial neural networks." *Journal of Construction Engineering and Management*, 146(3), 04019114.

Badri, A., Nadeau, S., and Gbodossou, A. (2012). "Proposal of a risk-factor-based analytical approach for integrating occupational health and safety into project risk evaluation." *Accident Analysis and Prevention*, Elsevier Ltd, 48, 223–234.

Barber, P., Graves, A., Hall, M., Sheath, D., and Tomkins, C. (2000). "Quality failure costs in civil engineering projects." *International Journal of Quality and Reliability Management*, 17(4), 479–492.

Chen, F., Jiao, H., Han, L., Shen, L., Du, W., Ye, Q., and Yu, G. (2020). "Real-time monitoring of construction quality for gravel piles based on Internet of Things." *Automation in Construction*, Elsevier, 116, 103228.

Chen, J. H., and Hsu, S. C. (2007). "Hybrid ANN-CBR model for disputed change orders in construction projects." *Automation in Construction*, 17(1), 56–64.

Cheng, M. Y., Tsai, H. C., and Chiu, Y. H. (2009). "Fuzzy case-based reasoning for coping with construction disputes." *Expert Systems with Applications*, Elsevier Ltd, 36, 4106–4113.

Choi, S., Kim, D. Y., Han, S. H., and Kwak, Y. H. (2014). "Conceptual cost-prediction model for public road planning via rough set theory and case-based reasoning." *Journal of Construction Engineering and Management*, 140(1), 04013026.

Doğan, S. Z., Arditi, D., and Günaydın, H. M. (2006). "Determining attribute weights in a CBR model for early cost prediction of structural systems." *Journal of Construction Engineering and Management*, 132(10), 1092–1098.

Doğan, S. Z., Arditi, D., and Murat Günaydin, H. (2008). "Using decision trees for determining attribute weights in a case-based model of early cost prediction." *Journal of Construction Engineering and Management*, 134(2), 146–152.

Dzeng, R.-J., and Tommelein, I. D. (1997). "Boiler erection scheduling using product models and case-based reasoning." *Journal of Construction Engineering and Management*, 123(3), 338–347.

Dzeng, R. J., and Tommelein, I. D. (2004). "Product modeling to support case-based construction planning and scheduling." *Automation in Construction*, 13(3), 341–360.

Fan, Z. P., Li, Y. H., and Zhang, Y. (2015). "Generating project risk response strategies based on CBR: A case study." *Expert Systems with Applications*, Elsevier Ltd, 42(6), 2870–2883.

Forbes, D., Smith, S. D., and Horner, M. (2007). "A case-based reasoning approach for selecting risk management techniques." *Association of Researchers in Construction Management, ARCOM 2007 - Proceedings of the 23rd Annual Conference*, 735–744.

Forcada, N., Gangolells, M., Casals, M., and Macarulla, M. (2017). "Factors Affecting Rework Costs in Construction." *Journal of Construction Engineering and Management*, 143(8), 04017032.

Hallowell, M. R., and Gambatese, J. A. (2010). "Qualitative Research: Application of the Delphi Method to CEM Research." *Journal of Construction Engineering and Management*, 136(1), 99–107.

Hong, T., Hyun, C., and Moon, H. (2011). "CBR-based cost prediction model-II of the design phase for multi-family housing projects." *Expert Systems with Applications*, Elsevier Ltd, 38(3), 2797–2808.

Hu, X., Xia, B., Skitmore, M., and Chen, Q. (2016). "The application of case-based reasoning in construction management research: An overview." *Automation in Construction*, Elsevier B.V., 72, 65–74.

Hwang, B.-G., Thomas, S. R., Haas, C. T., and Caldas, C. H. (2009). "Measuring the Impact of Rework on Construction Cost Performance." *Journal of Construction Engineering and Management*, 135(3), 187–198.

Jafari, A., and Love, P. E. D. (2013). "Quality Costs in construction: Case of Qom Monorail Project in Iran." *Journal of Construction Engineering and Management*, 139(9), 1244–1249.

Ji, C., Hong, T., and Hyun, C. (2010). "CBR revision model for ımproving cost prediction accuracy in multifamily housing projects." *Journal of Management in Engineering*, 26(4), 229–236.

Ji, S.-H., Park, M., Lee, H.-S., Ahn, J., Kim, N., and Son, B. (2011). "Military facility cost estimation system using case-based reasoning in Korea." *Journal of Computing in Civil Engineering*, 25(3), 218–231.

Jin, R., Cho, K., Hyun, C., and Son, M. (2012). "MRA-based revised CBR model for cost prediction in the early stage of construction projects." *Expert Systems with Applications*, Elsevier Ltd, 39(5), 5214–5222.

Jin, R., Han, S., Hyun, C., and Kim, J. (2014). “Improving accuracy of early stage cost estimation by revising categorical variables in a case-based reasoning model.” *Journal of Construction Engineering and Management*, 140(7), 04014025.

Josephson, P.-E., and Hammarlund, Y. (1999). “The causes and costs of defects in construction.” *Automation in Construction*, 8, 681–687.

Kazemian, A., Yuan, X., Davtalab, O., and Khoshnevis, B. (2019). “Computer vision for real-time extrusion quality monitoring and control in robotic construction.” *Automation in Construction*, Elsevier, 101, 92–98.

Kim, B., and Hong, T. (2012). “Revised case-based reasoning model development based on multiple regression analysis for railroad bridge construction.” *Journal of Construction Engineering and Management*, 138(1), 154–162.

Kim, G. H., An, S. H., and Kang, K. I. (2004). “Comparison of construction cost estimating models based on regression analysis, neural networks, and case-based reasoning.” *Building and Environment*, 39(10), 1235–1242.

Kim, K. J., and Kim, K. (2010). “Preliminary Cost Estimation Model Using Case-Based Reasoning and Genetic Algorithms.” *Journal of Computing in Civil Engineering*, 24(6), 499–505.

Kim, S. (2012). “Interval estimation of construction cost using case-based reasoning and genetic algorithms.” *Journal of Asian Architecture and Building Engineering*, 11(2), 327–334.

Kim, S. Y., Choi, J. W., Kim, G. H., and Kang, K. I. (2005). “Comparing cost prediction methods for apartment housing projects: CBR versus ANN.” *Journal of Asian Architecture and Building Engineering*, 4(1), 113–120.

Koo, C. W., Hong, T. H., Hyun, C. T., and Koo, K. J. (2010). “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*, 37(5), 739–752.

Kuzucuoglu, D., Ayhan, B. U., Tokdemir, O. B. (2019). “Integrating lean construction into the civil engineering curriculum.” *The International Journal of Engineering Education*, 35(6A), 1787–1802.

Lee, S., Jin, Y., Woo, S., and Shin, D. H. (2013). “Approximate cost estimating model of eco-type trade for river facility construction using case-based reasoning and genetic algorithms.” *KSCE Journal of Civil Engineering*, 17(2), 292–300.

Li, H. (1996). “Case-based reasoning for intelligent support of construction negotiation.” *Information and Management*, 30(5), 231–238.

Love, P. E. D. (2002). “Influence of project type and procurement method on rework costs in building construction projects.” *Journal of Construction Engineering and Management*, 128(1), 18–29.

Love, P. E. D., and Edwards, D. J. (2004).” Forensic project management: The underlying causes of rework in construction projects.” *Civil Enginnering and Environmental Systems*, 21(3), 207-228.

Love, P. E. D. and Li, H. (2000a). “Quantifying the causes and costs of rework in construction”, *Construction Management and Economics*, 18(4), 479-90.

Love, P. E. D. and Li, H. (2000b). “Overcoming the problems associated with quality certification”, *Construction Management and Economics*, 18(2), 139-49.

Love, P. E. D., and Sing, C. P. (2013). “Determining the probability distribution of rework costs in construction and engineering projects.” *Structure and Infrastructure Engineering*, 9(11), 1136–1148.

Love, P. E. D., Smith, J., Ackermann, F., Irani, Z., and Teo, P. (2018). “The costs of rework: insights from construction and opportunities for learning.” *Production Planning and Control*, Taylor & Francis, 29(13), 1082–1095.

Love, P. E. D., Teo, P., and Morrison, J. (2017). “Revisiting quality failure costs in construction.” *Journal of Construction Engineering and Management*, 144(2), 05017020.

Marzouk, M. M., and Ahmed, R. M. (2011). “A case-based reasoning approach for estimating the costs of pump station projects.” *Journal of Advanced Research*, Cairo University, 2(4), 289–295.

Oke, A. E., and Ugoje, F. F. (2013). “Assessment of rework cost of selected building projects in Nigeria.” *International Journal of Quality and Reliability Management*, 30(7), 799–810.

Olayinka, R., Suresh, S., and Chinyio, E. (2016). “Impact of knowledge management on the cost of poor quality.” *Proceedings of the Institution of Civil Engineers - Management, Procurement and Law*, 177–188.

Ozorhon, B., Dikmen, I., and Birgonul, M. T. (2006). “Case-based reasoning model for ınternational market selection.” *Journal of Construction Engineering and Management*, 132(9), 940–948.

Palisade (2020). “Evolver 8.1: Sophisticated optimization for Excel” Accessed January 2021. https://www.palisade.com/evolver/

Ryu, H.-G., Lee, H.-S., and Park, M. (2007). “Construction planning method using case-based reasoning (CONPLA-CBR).” *Journal of Computing in Civil Engineering*, 21(6), 410–422.

Saaty, T. L. (1990). “How to make a decision: The analytic hierarchy process.” *European Journal of Operational Research*, 48(1), 9–26.

Saaty, T. L. (2008). “Decision making with the analytic hierarchy process.” *International Journal of Services Sciences*, 1(1), 83.

Seyis, S., and Ergen, E. (2017). “A decision making support tool for selecting green building certification credits based on project delivery attributes.” *Building and Environment*, Elsevier Ltd, 126, 107–118.

Sheng, D., Ding, L., Zhong, B., Love, P. E. D., Luo, H., and Chen, J. (2020). “Construction quality information management with blockchains.” *Automation in Construction*, Elsevier, 120, 103373.

Zhong, D., Li, X., Cui, B., Wu, B., and Liu, Y. (2018). “Technology and application of real-time compaction quality monitoring for earth-rockfill dam construction in deep narrow valley.” *Automation in Construction*, Elsevier, 90, 23–38.

APPENDICES

APPENDIX-A

Table A.1. *Actual Cost Impacts for Test Cases*

Test Case ID	Cost Impacts				
	1	2	3	4	5
18	20,00%	50,00%	20,00%	10,00%	0,00%
21	20,00%	55,00%	15,00%	10,00%	0,00%
44	30,00%	53,33%	10,00%	6,67%	0,00%
45	37,50%	50,00%	7,50%	5,00%	0,00%
47	40,00%	48,00%	8,00%	4,00%	0,00%
49	43,33%	45,00%	8,33%	3,33%	0,00%
50	45,71%	42,86%	8,57%	2,86%	0,00%
53	46,25%	42,50%	8,75%	2,50%	0,00%
66	45,56%	41,11%	10,00%	2,22%	1,11%
109	45,00%	39,00%	10,00%	4,00%	2,00%
122	42,73%	41,82%	9,09%	4,55%	1,82%
146	41,67%	43,33%	9,17%	4,17%	1,67%
152	39,23%	43,85%	10,77%	4,62%	1,54%
157	37,14%	43,57%	12,14%	5,71%	1,43%
173	34,67%	41,33%	14,00%	8,67%	1,33%
220	34,38%	40,00%	14,38%	10,00%	1,25%
238	35,29%	39,41%	14,12%	10,00%	1,18%
255	35,56%	40,00%	13,33%	10,00%	1,11%
264	35,26%	38,95%	14,21%	10,53%	1,05%
268	35,00%	38,00%	15,00%	11,00%	1,00%
275	33,81%	39,52%	14,76%	10,95%	0,95%
300	33,64%	40,45%	14,09%	10,91%	0,91%
326	33,04%	40,00%	15,22%	10,87%	0,87%
359	32,92%	39,58%	16,25%	10,42%	0,83%
394	32,40%	40,80%	16,00%	10,00%	0,80%
396	31,92%	41,54%	16,15%	9,62%	0,77%
403	31,85%	41,85%	15,93%	9,26%	1,11%
446	32,14%	41,43%	15,71%	9,64%	1,07%
456	32,41%	41,72%	15,17%	9,66%	1,03%
457	32,67%	42,00%	14,67%	9,67%	1,00%
458	32,58%	42,26%	14,19%	10,00%	0,97%
468	32,50%	42,19%	14,38%	10,00%	0,94%
476	33,03%	41,82%	14,55%	9,70%	0,91%
510	33,24%	42,35%	14,12%	9,41%	0,88%
512	34,00%	42,29%	13,71%	9,14%	0,86%

Table A.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
513	34,72%	42,22%	13,33%	8,89%	0,83%
541	34,32%	42,97%	13,24%	8,65%	0,81%
569	34,47%	43,42%	12,89%	8,42%	0,79%
581	34,87%	43,59%	12,56%	8,21%	0,77%
615	36,25%	42,75%	12,25%	8,00%	0,75%
640	36,83%	42,68%	11,95%	7,80%	0,73%
641	37,14%	42,62%	11,90%	7,62%	0,71%
661	37,67%	42,33%	11,86%	7,44%	0,70%
670	38,41%	41,82%	11,82%	7,27%	0,68%
718	39,33%	41,11%	11,56%	7,33%	0,67%
727	40,00%	40,65%	11,30%	7,39%	0,65%
760	40,43%	40,43%	11,28%	7,23%	0,64%
773	40,83%	40,21%	11,25%	7,08%	0,63%
784	41,02%	39,80%	11,63%	6,94%	0,61%
792	41,60%	39,40%	11,60%	6,80%	0,60%
818	42,55%	38,82%	11,37%	6,67%	0,59%
856	42,88%	38,46%	11,54%	6,54%	0,58%
878	43,21%	38,30%	11,51%	6,42%	0,57%
884	43,70%	37,78%	11,48%	6,48%	0,56%
921	44,36%	37,45%	11,27%	6,36%	0,55%
923	45,00%	37,14%	11,07%	6,25%	0,54%
926	45,61%	36,67%	11,05%	6,14%	0,53%
933	46,03%	36,38%	11,03%	6,03%	0,52%
979	46,10%	36,10%	11,36%	5,93%	0,51%
990	46,50%	36,00%	11,17%	5,83%	0,50%
1001	46,39%	35,90%	11,48%	5,74%	0,49%
1010	46,77%	35,48%	11,45%	5,81%	0,48%
1019	46,83%	35,56%	11,43%	5,71%	0,48%
1069	47,66%	35,00%	11,25%	5,63%	0,47%
1084	48,31%	34,62%	11,08%	5,54%	0,46%
1103	48,64%	34,55%	10,91%	5,45%	0,45%
1135	49,25%	34,18%	10,75%	5,37%	0,45%
1151	49,71%	33,97%	10,59%	5,29%	0,44%
1155	50,14%	33,62%	10,58%	5,22%	0,43%
1156	50,57%	33,29%	10,57%	5,14%	0,43%

Table A.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
1170	50,99%	33,10%	10,42%	5,07%	0,42%
1200	51,11%	33,06%	10,42%	5,00%	0,42%
1203	51,23%	33,01%	10,41%	4,93%	0,41%
1221	51,22%	32,97%	10,54%	4,86%	0,41%
1225	51,20%	33,07%	10,53%	4,80%	0,40%
1259	51,58%	32,89%	10,39%	4,74%	0,39%
1272	52,08%	32,60%	10,26%	4,68%	0,39%
1276	52,44%	32,31%	10,26%	4,62%	0,38%
1294	52,78%	32,03%	10,25%	4,56%	0,38%
1300	52,63%	32,13%	10,25%	4,63%	0,38%
1310	52,72%	32,22%	10,12%	4,57%	0,37%
1327	52,80%	32,07%	10,12%	4,63%	0,37%
1345	53,25%	31,81%	10,00%	4,58%	0,36%
1353	53,69%	31,55%	9,88%	4,52%	0,36%
1434	54,00%	31,29%	9,88%	4,47%	0,35%
1437	54,19%	31,16%	9,88%	4,42%	0,35%
1438	54,25%	31,15%	9,89%	4,37%	0,34%
1447	54,32%	31,14%	9,77%	4,43%	0,34%
1469	54,38%	31,12%	9,78%	4,38%	0,34%
1470	54,56%	31,00%	9,78%	4,33%	0,33%
1499	54,73%	30,99%	9,67%	4,29%	0,33%
1500	54,89%	30,98%	9,57%	4,24%	0,33%
1512	54,95%	30,97%	9,57%	4,19%	0,32%
1526	55,11%	30,85%	9,57%	4,15%	0,32%
1547	55,47%	30,53%	9,58%	4,11%	0,32%
1585	55,73%	30,42%	9,48%	4,06%	0,31%
1626	55,77%	30,52%	9,38%	4,02%	0,31%
1674	56,02%	30,41%	9,29%	3,98%	0,31%
1681	56,16%	30,40%	9,19%	3,94%	0,30%
1687	56,20%	30,50%	9,10%	3,90%	0,30%
1696	56,14%	30,59%	9,11%	3,86%	0,30%
1704	56,08%	30,78%	9,02%	3,82%	0,29%
1706	56,12%	30,87%	8,93%	3,79%	0,29%
1714	56,35%	30,77%	8,85%	3,75%	0,29%
1720	56,67%	30,57%	8,76%	3,71%	0,29%

Table A.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
1729	56,98%	30,38%	8,68%	3,68%	0,28%
1750	57,10%	30,28%	8,69%	3,64%	0,28%
1776	57,50%	30,00%	8,61%	3,61%	0,28%
1795	57,43%	30,09%	8,62%	3,58%	0,28%
1813	57,27%	30,09%	8,82%	3,55%	0,27%
1837	57,30%	30,00%	8,83%	3,60%	0,27%
1841	57,41%	30,00%	8,75%	3,57%	0,27%
1843	57,43%	30,09%	8,67%	3,54%	0,27%
1858	57,54%	30,09%	8,60%	3,51%	0,26%
1905	57,74%	29,91%	8,61%	3,48%	0,26%
1909	58,02%	29,66%	8,62%	3,45%	0,26%
1933	57,95%	29,83%	8,55%	3,42%	0,26%
1947	58,05%	29,75%	8,56%	3,39%	0,25%
1965	58,07%	29,83%	8,49%	3,36%	0,25%
1969	58,33%	29,67%	8,42%	3,33%	0,25%
2011	58,43%	29,59%	8,43%	3,31%	0,25%
2036	58,52%	29,51%	8,44%	3,28%	0,25%
2040	58,62%	29,51%	8,37%	3,25%	0,24%
2066	58,71%	29,44%	8,31%	3,31%	0,24%
2079	58,80%	29,28%	8,32%	3,28%	0,32%
2081	58,89%	29,13%	8,33%	3,25%	0,40%
2083	58,90%	28,98%	8,35%	3,23%	0,55%
2118	59,06%	28,91%	8,28%	3,20%	0,55%
2147	59,15%	28,91%	8,22%	3,18%	0,54%
2149	59,23%	28,92%	8,15%	3,15%	0,54%
2153	59,39%	28,85%	8,09%	3,13%	0,53%
2164	59,39%	28,79%	8,11%	3,18%	0,53%
2177	59,47%	28,80%	8,05%	3,16%	0,53%
2185	59,40%	28,88%	8,06%	3,13%	0,52%
2197	59,33%	28,96%	8,00%	3,19%	0,52%
2218	59,56%	28,75%	8,01%	3,16%	0,51%
2237	59,64%	28,69%	8,03%	3,14%	0,51%
2239	59,71%	28,62%	8,04%	3,12%	0,51%
2240	59,78%	28,56%	8,06%	3,09%	0,50%
2251	59,86%	28,57%	8,00%	3,07%	0,50%

Table A.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
2268	59,86%	28,51%	8,01%	3,12%	0,50%
2299	60,07%	28,38%	7,96%	3,10%	0,49%
2301	60,21%	28,25%	7,97%	3,08%	0,49%
2343	60,35%	28,06%	8,06%	3,06%	0,49%
2381	60,34%	28,14%	8,00%	3,03%	0,48%
2404	60,34%	28,08%	8,01%	3,08%	0,48%
2448	60,34%	28,03%	8,10%	3,06%	0,48%
2467	60,47%	27,91%	8,11%	3,04%	0,47%
2481	60,40%	28,05%	8,05%	3,02%	0,47%
2497	60,40%	28,07%	8,07%	3,00%	0,47%

APPENDIX-B

Table B.1. *Predicted Cost Impacts with CBR-AHP Model for Test Cases*

Test Case ID	Cost Impacts				
	1	2	3	4	5
18	50,31%	28,77%	10,92%	8,46%	1,54%
21	65,08%	25,69%	6,92%	2,31%	0,00%
44	62,78%	25,14%	7,40%	3,78%	0,90%
45	67,34%	23,34%	6,16%	2,54%	0,62%
47	67,44%	23,04%	6,76%	2,54%	0,22%
49	67,32%	23,18%	6,74%	2,56%	0,20%
50	69,54%	22,32%	6,16%	1,86%	0,12%
53	66,54%	22,70%	7,28%	3,10%	0,38%
66	68,78%	22,62%	6,26%	2,18%	0,16%
109	68,98%	22,56%	6,30%	2,08%	0,08%
122	69,00%	22,58%	6,28%	2,00%	0,14%
146	68,70%	22,78%	6,16%	2,12%	0,24%
152	68,40%	22,82%	6,18%	2,26%	0,34%
157	69,22%	22,60%	6,22%	1,90%	0,06%
173	67,06%	22,68%	7,10%	2,84%	0,32%
220	68,80%	22,64%	6,26%	2,12%	0,18%
238	70,32%	22,40%	5,98%	1,20%	0,10%
255	71,10%	21,86%	5,90%	1,02%	0,12%
264	71,14%	21,80%	5,92%	1,04%	0,10%
268	71,24%	21,50%	6,00%	1,08%	0,18%
275	71,16%	21,82%	5,88%	1,02%	0,12%
300	70,80%	22,08%	5,86%	1,16%	0,10%
326	62,84%	25,18%	7,32%	3,76%	0,90%
359	65,88%	24,50%	6,92%	2,42%	0,28%
394	66,64%	23,52%	6,64%	2,88%	0,32%
396	62,75%	25,19%	7,39%	3,77%	0,90%
403	68,86%	22,87%	6,15%	2,06%	0,06%
446	69,34%	22,38%	6,09%	2,06%	0,14%
456	70,32%	22,14%	5,77%	1,70%	0,08%
457	68,26%	22,67%	6,79%	2,12%	0,16%
458	62,75%	25,19%	7,39%	3,77%	0,90%
468	62,99%	24,97%	7,54%	3,69%	0,80%
476	63,81%	24,97%	7,78%	3,05%	0,38%
510	64,03%	24,87%	7,68%	2,93%	0,48%
512	65,63%	24,33%	7,29%	2,46%	0,30%

Table B.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
513	66,29%	24,07%	7,15%	2,26%	0,24%
541	70,92%	21,42%	6,25%	1,36%	0,06%
569	71,32%	21,64%	5,87%	1,12%	0,06%
581	70,58%	22,34%	5,45%	1,40%	0,24%
615	71,86%	21,52%	5,51%	0,94%	0,18%
640	69,80%	22,00%	6,45%	1,54%	0,22%
641	71,72%	21,46%	5,63%	0,98%	0,22%
661	71,56%	21,40%	5,75%	1,04%	0,26%
670	71,70%	22,04%	5,31%	0,86%	0,10%
718	71,08%	22,00%	5,61%	1,14%	0,18%
727	71,24%	21,64%	5,81%	1,26%	0,06%
760	71,24%	21,66%	5,81%	1,24%	0,06%
773	71,24%	21,66%	5,81%	1,24%	0,06%
784	71,42%	21,54%	5,75%	1,24%	0,06%
792	62,75%	25,19%	7,39%	3,77%	0,90%
818	71,08%	21,76%	5,77%	1,22%	0,18%
856	70,34%	21,50%	6,35%	1,62%	0,20%
878	71,48%	21,32%	5,89%	1,18%	0,14%
884	72,14%	21,10%	5,65%	1,02%	0,10%
921	71,24%	21,40%	5,97%	1,34%	0,06%
923	71,24%	21,40%	5,97%	1,34%	0,06%
926	62,75%	25,19%	7,39%	3,77%	0,90%
933	67,41%	23,07%	6,77%	2,53%	0,22%
979	67,47%	23,05%	6,79%	2,50%	0,20%
990	68,68%	23,19%	6,05%	2,00%	0,08%
1001	69,10%	22,79%	6,09%	1,86%	0,16%
1010	68,96%	22,97%	6,01%	1,88%	0,18%
1019	70,40%	21,90%	6,01%	1,58%	0,12%
1069	71,88%	21,14%	5,51%	1,38%	0,10%
1084	72,32%	20,38%	5,71%	1,46%	0,14%
1103	62,75%	25,19%	7,39%	3,77%	0,90%
1135	71,08%	21,76%	5,77%	1,22%	0,18%
1151	71,22%	21,24%	5,97%	1,36%	0,22%
1155	70,84%	21,96%	5,87%	1,20%	0,14%
1156	70,70%	21,72%	6,01%	1,36%	0,22%

Table B.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
1170	71,66%	21,30%	5,81%	1,02%	0,22%
1200	62,75%	25,19%	7,39%	3,77%	0,90%
1203	62,75%	25,19%	7,39%	3,77%	0,90%
1221	62,75%	25,19%	7,39%	3,77%	0,90%
1225	67,41%	23,07%	6,77%	2,53%	0,22%
1259	68,00%	22,79%	6,61%	2,40%	0,20%
1272	62,75%	25,19%	7,39%	3,77%	0,90%
1276	62,75%	25,19%	7,39%	3,77%	0,90%
1294	69,34%	22,38%	6,09%	2,06%	0,14%
1300	69,34%	22,38%	6,09%	2,06%	0,14%
1310	70,08%	22,14%	5,93%	1,82%	0,04%
1327	70,18%	22,02%	5,89%	1,82%	0,10%
1345	70,14%	22,10%	6,23%	1,50%	0,04%
1353	62,75%	25,19%	7,39%	3,77%	0,90%
1434	62,75%	25,19%	7,39%	3,77%	0,90%
1437	62,75%	25,19%	7,39%	3,77%	0,90%
1438	62,75%	25,19%	7,39%	3,77%	0,90%
1447	63,19%	25,09%	7,31%	3,57%	0,84%
1469	67,41%	23,07%	6,77%	2,53%	0,22%
1470	67,64%	23,01%	6,81%	2,38%	0,16%
1499	68,72%	22,63%	6,29%	2,18%	0,18%
1500	68,72%	22,63%	6,29%	2,18%	0,18%
1512	68,72%	23,81%	5,69%	1,52%	0,26%
1526	72,51%	21,20%	4,99%	1,16%	0,14%
1547	72,50%	21,16%	5,03%	1,20%	0,12%
1585	62,75%	25,19%	7,39%	3,77%	0,90%
1626	69,50%	22,50%	5,93%	1,88%	0,20%
1674	69,10%	22,55%	6,03%	2,18%	0,14%
1681	70,40%	22,04%	6,17%	1,28%	0,12%
1687	62,75%	25,19%	7,39%	3,77%	0,90%
1696	63,61%	25,17%	7,33%	3,31%	0,58%
1704	63,13%	25,47%	7,39%	3,43%	0,58%
1706	70,86%	22,14%	5,31%	1,46%	0,24%
1714	71,50%	21,38%	5,33%	1,60%	0,20%
1720	72,34%	21,02%	5,13%	1,36%	0,16%

Table B.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
1729	62,75%	25,19%	7,39%	3,77%	0,90%
1750	71,10%	21,84%	5,95%	1,00%	0,12%
1776	71,06%	21,90%	5,91%	1,02%	0,12%
1795	70,40%	22,34%	5,93%	1,20%	0,14%
1813	70,54%	21,88%	6,15%	1,34%	0,10%
1837	71,38%	21,46%	5,93%	1,14%	0,10%
1841	71,10%	21,84%	5,95%	1,00%	0,12%
1843	70,54%	21,88%	6,15%	1,34%	0,10%
1858	71,10%	21,84%	5,95%	1,00%	0,12%
1905	70,72%	22,02%	5,99%	1,16%	0,12%
1909	71,08%	21,76%	5,77%	1,22%	0,18%
1933	72,18%	21,62%	5,23%	0,90%	0,08%
1947	70,86%	22,14%	5,31%	1,46%	0,24%
1965	72,53%	21,54%	4,87%	0,96%	0,10%
1969	72,51%	21,00%	5,19%	1,08%	0,22%
2011	71,50%	21,38%	5,33%	1,60%	0,20%
2036	70,86%	22,14%	5,31%	1,46%	0,24%
2040	71,00%	22,10%	5,25%	1,42%	0,24%
2066	71,68%	21,70%	5,27%	1,24%	0,12%
2079	72,38%	21,14%	5,25%	1,06%	0,18%
2081	72,20%	21,00%	5,23%	1,42%	0,16%
2083	72,59%	20,90%	5,19%	1,10%	0,22%
2118	72,55%	20,96%	5,13%	1,10%	0,26%
2147	71,26%	22,02%	5,19%	1,36%	0,18%
2149	71,26%	22,02%	5,19%	1,36%	0,18%
2153	71,36%	21,60%	5,57%	1,30%	0,18%
2164	70,70%	22,14%	5,71%	1,28%	0,18%
2177	72,26%	21,00%	5,39%	1,08%	0,28%
2185	70,86%	22,14%	5,31%	1,46%	0,24%
2197	62,75%	25,19%	7,39%	3,77%	0,90%
2218	70,86%	22,14%	5,31%	1,46%	0,24%
2237	71,22%	21,56%	5,37%	1,52%	0,34%
2239	72,97%	20,62%	4,95%	1,28%	0,18%
2240	72,57%	21,04%	4,99%	1,24%	0,16%
2251	71,56%	21,44%	5,49%	1,34%	0,18%

Table B.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
2268	72,20%	21,22%	5,27%	1,12%	0,20%
2299	62,75%	25,19%	7,39%	3,77%	0,90%
2301	62,75%	25,19%	7,39%	3,77%	0,90%
2343	66,91%	23,45%	6,63%	2,73%	0,28%
2381	69,34%	22,38%	6,09%	2,06%	0,14%
2404	62,75%	25,19%	7,39%	3,77%	0,90%
2448	71,10%	21,84%	5,95%	1,00%	0,12%
2467	71,10%	21,84%	5,95%	1,00%	0,12%
2481	71,36%	21,44%	5,87%	1,18%	0,16%
2497	70,72%	22,06%	6,03%	1,06%	0,14%

APPENDIX-C

Table C.1. *Predicted Cost Impacts with CBR-AHP Model for Test Cases*

Test Case ID	Cost Impacts				
	1	2	3	4	5
18	64,06%	24,52%	7,03%	3,68%	0,71%
21	66,11%	24,77%	6,61%	1,97%	0,54%
44	61,57%	25,57%	8,05%	3,87%	0,94%
45	64,72%	24,93%	7,28%	2,53%	0,54%
47	61,57%	25,58%	8,04%	3,87%	0,94%
49	64,38%	25,17%	7,57%	2,49%	0,40%
50	64,46%	25,12%	7,56%	2,46%	0,40%
53	65,33%	24,75%	7,38%	2,18%	0,36%
66	68,84%	23,08%	6,48%	1,40%	0,21%
109	69,49%	22,75%	6,40%	1,23%	0,14%
122	69,42%	22,73%	6,26%	1,33%	0,26%
146	68,85%	23,04%	6,49%	1,40%	0,22%
152	68,78%	22,85%	6,37%	1,60%	0,40%
157	69,61%	22,67%	6,40%	1,17%	0,15%
173	68,54%	22,36%	7,03%	1,82%	0,25%
220	68,85%	23,09%	6,48%	1,35%	0,22%
238	69,43%	22,24%	6,12%	1,91%	0,30%
255	70,12%	22,31%	5,98%	1,34%	0,25%
264	70,40%	22,28%	5,93%	1,23%	0,17%
268	71,40%	21,40%	5,86%	1,13%	0,22%
275	70,17%	22,28%	5,97%	1,34%	0,25%
300	69,46%	22,20%	6,13%	1,91%	0,30%
326	61,62%	25,59%	8,00%	3,85%	0,94%
359	63,77%	25,03%	8,01%	2,76%	0,43%
394	69,48%	22,18%	6,12%	1,92%	0,30%
396	61,56%	25,60%	8,04%	3,86%	0,94%
403	67,35%	24,11%	6,15%	2,11%	0,28%
446	67,35%	24,11%	6,15%	2,11%	0,28%
456	67,67%	23,90%	6,10%	2,06%	0,28%
457	68,83%	23,48%	5,90%	1,61%	0,18%
458	61,56%	25,60%	8,04%	3,86%	0,94%
468	61,56%	25,60%	8,04%	3,86%	0,94%
476	62,25%	25,48%	8,01%	3,46%	0,80%
510	68,80%	23,09%	6,50%	1,39%	0,22%
512	70,06%	22,14%	6,41%	1,20%	0,19%

Table C.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
513	70,57%	22,07%	6,18%	0,98%	0,21%
541	67,35%	24,11%	6,15%	2,11%	0,28%
569	68,83%	23,48%	5,90%	1,61%	0,18%
581	69,41%	22,61%	5,70%	1,92%	0,37%
615	61,56%	25,60%	8,04%	3,86%	0,94%
640	62,91%	25,56%	7,82%	3,14%	0,57%
641	68,52%	23,03%	6,50%	1,74%	0,21%
661	68,52%	23,03%	6,50%	1,74%	0,21%
670	69,19%	23,16%	6,06%	1,43%	0,17%
718	69,41%	22,25%	6,12%	1,92%	0,30%
727	67,35%	24,11%	6,15%	2,11%	0,28%
760	67,38%	24,08%	6,17%	2,10%	0,28%
773	72,03%	21,09%	5,37%	1,26%	0,26%
784	72,08%	21,05%	5,38%	1,23%	0,26%
792	61,56%	25,60%	8,04%	3,86%	0,94%
818	69,41%	22,25%	6,12%	1,92%	0,30%
856	71,10%	21,90%	5,64%	1,21%	0,14%
878	71,48%	21,31%	5,72%	1,27%	0,22%
884	61,56%	25,60%	8,04%	3,86%	0,94%
921	67,35%	24,11%	6,15%	2,11%	0,28%
923	67,35%	24,11%	6,15%	2,11%	0,28%
926	61,56%	25,60%	8,04%	3,86%	0,94%
933	61,56%	25,60%	8,04%	3,86%	0,94%
979	65,52%	25,03%	7,16%	2,04%	0,25%
990	69,41%	22,25%	6,12%	1,92%	0,30%
1001	72,12%	21,06%	5,54%	1,12%	0,15%
1010	61,56%	25,60%	8,04%	3,86%	0,94%
1019	67,35%	24,11%	6,15%	2,11%	0,28%
1069	61,96%	25,23%	8,14%	3,78%	0,90%
1084	65,32%	24,69%	7,46%	2,23%	0,29%
1103	61,56%	25,60%	8,04%	3,86%	0,94%
1135	69,41%	22,25%	6,12%	1,92%	0,30%
1151	70,08%	22,34%	5,97%	1,35%	0,25%
1155	69,41%	22,25%	6,12%	1,92%	0,30%
1156	70,33%	21,74%	5,88%	1,67%	0,39%

Table C.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
1170	72,29%	21,08%	5,42%	1,06%	0,15%
1200	61,56%	25,60%	8,04%	3,86%	0,94%
1203	61,56%	25,60%	8,04%	3,86%	0,94%
1221	61,56%	25,60%	8,04%	3,86%	0,94%
1225	61,56%	25,60%	8,04%	3,86%	0,94%
1259	61,96%	25,23%	8,14%	3,78%	0,90%
1272	61,56%	25,60%	8,04%	3,86%	0,94%
1276	61,56%	25,60%	8,04%	3,86%	0,94%
1294	67,35%	24,11%	6,15%	2,11%	0,28%
1300	67,35%	24,11%	6,15%	2,11%	0,28%
1310	70,33%	22,52%	5,38%	1,50%	0,26%
1327	70,33%	22,52%	5,38%	1,50%	0,26%
1345	69,41%	22,61%	5,70%	1,92%	0,37%
1353	61,56%	25,60%	8,04%	3,86%	0,94%
1434	61,56%	25,60%	8,04%	3,86%	0,94%
1437	61,56%	25,60%	8,04%	3,86%	0,94%
1438	61,56%	25,60%	8,04%	3,86%	0,94%
1447	61,56%	25,60%	8,04%	3,86%	0,94%
1469	61,56%	25,60%	8,04%	3,86%	0,94%
1470	61,86%	25,53%	7,94%	3,75%	0,91%
1499	68,80%	23,09%	6,50%	1,39%	0,22%
1500	68,80%	23,09%	6,50%	1,39%	0,22%
1512	67,35%	24,11%	6,15%	2,11%	0,28%
1526	70,03%	22,65%	5,60%	1,57%	0,15%
1547	70,03%	22,65%	5,60%	1,57%	0,15%
1585	61,56%	25,60%	8,04%	3,86%	0,94%
1626	68,52%	23,03%	6,50%	1,74%	0,21%
1674	68,69%	23,09%	6,30%	1,70%	0,22%
1681	67,35%	24,11%	6,15%	2,11%	0,28%
1687	61,56%	25,60%	8,04%	3,86%	0,94%
1696	64,03%	25,09%	7,75%	2,70%	0,43%
1704	64,74%	24,92%	7,56%	2,40%	0,37%
1706	69,41%	22,25%	6,12%	1,92%	0,30%
1714	67,35%	24,11%	6,15%	2,11%	0,28%
1720	70,33%	22,52%	5,38%	1,50%	0,26%

Table C.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
1729	61,56%	25,60%	8,04%	3,86%	0,94%
1750	68,80%	23,09%	6,50%	1,39%	0,22%
1776	68,70%	22,88%	6,40%	1,61%	0,40%
1795	67,31%	24,10%	6,18%	2,14%	0,28%
1813	71,09%	22,17%	5,27%	1,28%	0,19%
1837	72,00%	21,67%	5,10%	1,08%	0,15%
1841	70,17%	22,47%	5,60%	1,48%	0,29%
1843	71,41%	21,94%	5,17%	1,24%	0,23%
1858	70,17%	22,47%	5,60%	1,48%	0,29%
1905	69,38%	22,61%	5,70%	1,94%	0,37%
1909	69,30%	22,65%	5,74%	1,94%	0,37%
1933	72,37%	21,66%	4,83%	0,98%	0,17%
1947	69,30%	22,65%	5,74%	1,94%	0,37%
1965	72,37%	21,66%	4,83%	0,98%	0,17%
1969	71,37%	21,71%	5,34%	1,30%	0,29%
2011	67,35%	24,11%	6,15%	2,11%	0,28%
2036	69,41%	22,61%	5,70%	1,92%	0,37%
2040	70,33%	22,10%	5,45%	1,67%	0,46%
2066	69,41%	22,61%	5,70%	1,92%	0,37%
2079	72,12%	21,42%	5,12%	1,12%	0,22%
2081	72,03%	21,52%	5,02%	1,21%	0,22%
2083	71,48%	21,67%	5,30%	1,27%	0,29%
2118	72,06%	21,46%	5,14%	1,09%	0,25%
2147	67,13%	24,19%	6,29%	2,11%	0,28%
2149	67,13%	24,19%	6,29%	2,11%	0,28%
2153	69,61%	22,91%	5,83%	1,49%	0,15%
2164	67,35%	24,11%	6,15%	2,11%	0,28%
2177	72,03%	21,09%	5,37%	1,26%	0,26%
2185	69,41%	22,61%	5,70%	1,92%	0,37%
2197	61,56%	25,60%	8,04%	3,86%	0,94%
2218	69,41%	22,25%	6,12%	1,92%	0,30%
2237	71,34%	21,39%	5,63%	1,39%	0,25%
2239	72,94%	20,50%	5,38%	1,06%	0,12%
2240	72,25%	21,09%	5,39%	1,13%	0,14%
2251	72,04%	21,09%	5,63%	1,12%	0,12%

Table C.1. *(Continued)*

Test Case ID	Cost Impacts				
	1	2	3	4	5
2268	72,25%	21,06%	5,43%	1,08%	0,18%
2299	61,56%	25,60%	8,04%	3,86%	0,94%
2301	61,56%	25,60%	8,04%	3,86%	0,94%
2343	69,41%	22,25%	6,12%	1,92%	0,30%
2381	67,35%	24,11%	6,15%	2,11%	0,28%
2404	61,56%	25,60%	8,04%	3,86%	0,94%
2448	68,80%	23,09%	6,50%	1,39%	0,22%
2467	68,80%	23,09%	6,50%	1,39%	0,22%
2481	71,06%	21,78%	6,03%	1,02%	0,11%
2497	61,56%	25,60%	8,04%	3,86%	0,94%