Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 22, Sayı 3, 1132-1142, 2018

Süleyman Demirel University
Journal of Natural and Applied Sciences
Volume 22, Issue 3, 1132-1142, 2018

DOI: 10.19113/sdufenbed.469288

Reference Class Forecasting Method in Predicting Construction Project Duration

Savaş BAYRAM*[1], Saad AL-JIBOURI[2]

[1]Erciyes University, Faculty of Engineering, Department of Civil Engineering, 38280, Kayseri, Turkey
[2]University of Twente, Faculty of Engineering Technology, Department of Construction Management&Engineering, P.O. Box 217, 7500 AE Enschede, The Netherlands

(Alınış / Received: 23.01.2018, Kabul / Accepted: 04.10.2018, Online Yayınlanma/ Published Online: 10.10.2018)

Keywords
Construction,
Project management,
Estimating,
Scheduling

Abstract: Project delay is a global problem affecting construction and other industries in many countries. Its impact on planning and budgeting can be serious for all stakeholders involved and difficult to resolve. The purpose of this study is to analyze the reliability of duration estimates of public building projects based on actual duration of similar projects carried out in the past.Turkey is used as a case study for this purpose and data from 643 public building projects completed in Turkey were collected. The data include contract durations, actual durations as well as the total construction areas for all projects. Reference Class Forecasting (RCF) method is proposed and used to investigate whether it would be possible to produce reliable and realistic project duration forecasts based on such data. RCF can realistically predict the actual final duration of the projects of different reference classes for various levels of acceptable risks. Original estimates of the contract durations in Turkey are generally optimistic or underestimated. Government buildings with the highest average construction area required lower uplift values on the estimated durations to produce accurate and realistic forecasts.So far the RCF method has been broadly applied to predict project cost rather than duration. This paper describes it use for forecasting duration in building projects.

İnşaat Proje Süresi Tahmininde Referans Sınıf Tahmin Yöntemi

Anahtar Kelimeler
Yapı,
Proje yönetimi,
Tahminleme,
İş programı

Özet: Projelerde yaşanan süresel gecikmeler, pek çok ülkede inşaat sektörünü ve diğer sektörleri etkileyen küresel bir sorundur. Süresel gecikmelerin planlama ve bütçeleme üzerindeki etkisi, ilgili tüm paydaşlar için ciddi ve çözülmesi zor olmaktadır. Bu çalışmanın amacı, geçmişte tamamlanan benzer projelerin gerçekleşen sürelerine dayalı olarak, kamu binası projelerinin süresel tahminlerinin güvenilirliğini analiz etmektir. Bu amaçla Türkiye'de tamamlanan 643 kamu inşaat projesinin verileri temin edilmiştir. Proje verileri; sözleşme süreleri, fiili gerçekleşme süreleri ve toplam inşaat alanlarından oluşmaktadır. Bu verilere dayalı olarak güvenilir ve gerçekçi proje süresi tahminleri üretmenin mümkün olup olmayacağını irdelemek için Referans Sınıf Tahmin (RST) yöntemi önerilmiş ve kullanılmıştır. RST, kabul edilebilir risklerin farklı seviyeleri için değişik referans sınıflarının fiili proje sürelerini gerçekçi bir şekilde tahmin edebilen bir yaklaşımdır. Sonuç olarak; Türkiye'de gerçekleştirilen kamu inşaat projelerinin sözleşme süresi öngörülerinin genellikle iyimser veya olması gerekenden düşük olduğu belirlenmiştir. Doğru ve gerçekçi tahminler üretmek adına, en yüksek ortalama inşaat alanına sahip olan devlet binalarının, öngörülen proje süreleri üzerinden nispeten daha düşük yükseltme değerlerine sahip oldukları tespit edilmiştir. RST yöntemi, önceki çalışmalarda proje süresinden ziyade yapım maliyetini tahmin etmek için yaygın bir şekilde kullanılmıştır. Bu çalışmada ise RST yönteminin bina projelerinin süresel tahmininde kullanılabilirliği esas alınmıştır.

1. Introduction

The competitive nature of the construction industry exerts pressure on contractors to keep project duration and cost as low as possible [1]. The contract duration set by the client is not always realistic and many projects often experience delay [2]. Project delay is a global problem affecting not only the construction industry but the overall economy of countries [3,4]. It usually involves multiple complex issues all of which are invariably of critical importance to the parties to the construction contract

* *Corresponding author: sbayram@erciyes.edu.tr*

[5]. The consequences of delay can be very serious and hard to resolve [6]. Delay can result in additional charges for the client, professional fees and income lost through late occupancy [3,7].

Experience from construction in other developing and transition countries has also indicated serious problems as consequences of delay of projects on the economy [8]. For example the influence on the overall economy of a developing country like Turkey, taken here as an example, where construction represents 30% of the 'gross national product', can have a profound negative impact on the country's overall development programme [9]. Realistic project duration forecasts are therefore vital not only for the projects and the construction sector but for the economy as a whole. Experiences however shows that reliable project duration forecasts are still troubled and public construction projects cannot be completed in time. For instance, in Gaziantep, an 18-classroom school project was completed in 2005 in 420 days, in contrast to its contract duration of 179 days. This represents a 135% delay from the original plan. Another example from the same year is of a small 6-classroom school project in Izmir. The project contract duration was 79 days but the project was actually completed in 412 days; duration overrun of 422%.

This study is an attempt to use Reference Class Forecasting (RCF) to forecast building project duration based on data collected from the Turkish constructions. Literature on the application of RCF in construction management is limited and has so far been generally focused on cost forecasting. A few studies in the application of RCF to forecast project duration had been reported. For instance; Moret and Einstein (2016) modeled the four main types of structures (tunnels, viaducts, cuts and embankments) in rail lines based on Portuguese high-speed rail projects [10]. The largest increases were observed in the tunnel construction cost (58%) and in the earthwork like cuts and embankments construction duration (94%). Flyvbjerg et al. (2016) established RCF models to forecast the costs and the durations of 25 completed roadwork projects in Hong Kong [11]. For time-to-completion forecasting, it was found that generally more than half of the project duration was spent before actual construction that is the pre-construction stage. Batselier and Vanhoucke (2016) also used the RCF technique to compare with the Monte Carlo simulation and earned value management (EVM) for both cost and time forecasting of 56 projects [12]. It was found that the RCF technique was the most user-friendly, as it does not require a great deal of detailed information or extensive calculations. In this study the reliability of RCF in producing realistic project duration forecasts will be investigated and discussed. The objective of the paper is to examine whether, similar to cost forecasting, RCF method can produce reliable and realistic project duration forecasts, based on historical data of completed public projects in similar reference classes. As such the paper will provide new contribution to scientific and practical knowledge in the construction management field. The differences between contract and actual durations were established by categorizing the types of the building projects as; educational, health service, government and security for use in RCF application.

2. Research Background and Literature Review

Even though the examples projects described in the introduction part seem to be extreme, the problem of project delay however is still real. For instance, a study by Arditi et al. (1985) on construction durations indicated that delay is not uncommon [13]. On the basis of data collected from 126 private projects and 258 public projects, the authors showed that the average delays in those projects were 34.60% and 43.65% respectively. A recent study, performed by Erdis (2013), evaluated the duration and cost variances of Turkish public building Projects [14]. It was found that the actual duration of 50% of the projects carried out under the previous State Procurement Law, nr. 2886, exceeded the original estimated duration whilst this was only 20% in the case of projects carried out under the current Public Procurement Law, nr. 4734.

The previous studies generally refer negative impact of delays on project success. Generally, the objectives of most of these studies were to analyze the causes of project delays as well as to classify and evaluate delays as seen above in the examples from the Turkish construction industry. Over the past decade, several studies have been performed to make sense of construction duration delay in a number of countries. For example Couto and Teixeira (2007) conducted a survey of 125 Portuguese construction stakeholders including contractor, client, and consultant to investigate the causes of delay [6]. In their study 118 causes were identified, analyzed and the most important cause of delay was found to be incomplete design. Sweis et al. (2008) investigated the causes of delay in residential projects in Jordan.Changing orders of the owner are presented to be the leading cause of construction delay [15]. In another study in Egypt by Abd El-Razek et al. (2008) found that the most important causes of delay are available finance to contractor during construction and delays in contractor's payment by owner [16]. Also Fugar and Agyakwah-Baah (2010) identified 32 possible causes of delay and subjected them to a field survey in Ghana and found that financial group factors (e.g. difficulty in accessing credit, fluctuation in prices) ranked highest among the major factors causing delay in construction projects [5].

The literature review described above indicates the diversity of factors and risks that can influence the work and cause delay. In most cases delay seem to be inevitable in construction. Realizing the inevitability of construction delay, owners and contractors often try to rely on various forecasting techniques to provide them with reliable indications of realistic projects' duration for the purpose of planning. Construction duration forecasting techniques used in the early stages of a building project are usually focused on predicting construction duration based on project cost and/or various physical characteristics such as floor area, type of facade etc. During construction, contractors have used 'earned value analysis' (EVA) to forecast the final cost of a project based on current performance and for eventually taking corrective actions [17]. Various other models and techniques have been developed over the years to facilitate project duration forecasting. For example Bromilow (1969; 1974) developed a model known as Bromilow's time-cost (BTC) model that uses duration as a function of cost to predict project duration [18,19]. Although the model was adapted by other researchers and used in several countries, the results of many applications of the model were reported as being unreliable, see for examples Odabasi (2009), Dursun & Stoy (2011), Ogunsemi & Jagboro (2006), Walker (1994) [20,21,22,23].

The relationship between the project characteristics and durations has also been examined by a number of researchers. For instance; Love et al. (2005) argued that gross floor area and number of floors provide better prediction for construction duration than estimated construction costs [24]. Stoy et al. (2007) identified construction speed drivers based on German building projects, and concluded that construction speed indicators can serve as the basis for early determination of the construction duration [25]. Dursun and Stoy (2012) reported that cost of construction projects and gross external floor area are the major variables related to the construction duration in Germany [26]. In recent research, Guerrero et al. (2014) used multiple regression analysis to develop a forecast model that allows estimating project duration of new builds using 168 building projects carried out in Spain [27]. Project type, gross floor area (GFA), the cost/GFA relationship and number of floors were used as predictor variables. It was found that GFA has greater influence than cost on project duration but both factors are necessary to achieve forecasts with higher accuracy. Gab-Allah et. al. (2015) used artificial neural network (ANN) model for predicting the expected construction duration in early stage using 130 building projects in Egypt [28]. Testing the validity of the model showed that the model has a good prediction capability with a maximum error of 14%. In another study, Leu and Liu (2016) have identified three prediction models based on the total cost for large, medium, and small industrial building construction projects in Taiwan using back-propagation neural network (BP-NN), a type of ANN and the results have demonstrated a considerable applicability of the proposed methodology [29]. Jarkas (2016) collected data comprising construction area and number of floors above and below ground from 113 residential and 74 office buildings projects completed during the period 2004–2010 in the State of Kuwait and subsequently analyzed using BTC and multiple linear regression models [30]. Another prediction model based on a new neuro-fuzzy algorithm for estimating duration of construction projects was proposed by Vahdani et al. (2016) [31]. The proposed proposed neuro-fuzzy model showed a better generalization performance. Jin et al. (2016) developed a successful case-based reasoning (CBR) model in the preliminary stage using 83 multihousing projects [32]. An overview of studies found in the literature, as described above, are also provided in Table 1.

The literature review indicates that factors responsible for delay in construction projects may vary from country to country and from one situation to another. It also indicated that the effectiveness of the techniques used to produce reliable forecasts will also vary accordingly. Hence simple and practical approach to account for the risk of delay resulting from the combination of these factors, in a particular environment, is required.

In cost forecasting and due to inaccuracies associated with cost estimates produced by traditional methods, Flyvbjerg and others, see for examples Flyvbjerg (2004), Flyvbjerg et al. (2005), Flyvbjerg (2007), Salling and Leleur (2012), Eythorsdottir (2012) [33,34,35,36,37], and currently Flyvbjerg et al. (2016), Moret and Einstein (2016) [10,11] suggested the use of Reference Class Forecasting (RCF) to forecast the final cost of infrastructure and transportation projects to produce more realistic estimates. They argued that in conventional estimation methods the reason for inaccurate estimates leading to higher actual costs cannot be traced to technical explanations but to human errors and/or because promoters deliberately make optimistic forecasts to get projects approved. The tendency of construction projects to overrun was referred to as "optimism bias".

Literature on the application of RCF in construction management is very limited and has so far been generally focused on cost forecasting. This study however is an attempt to use RCF to forecast building project duration rather than building cost.

3. Methodology

"Reference Class Forecasting" (RCF) is not a detailed estimation method but a forecast tool based on the work of Princeton psychologist Daniel Kahneman who won the Nobel prize in economics in 2002 [38]. Kahneman (1994) argued that the fallacy originates from actors taking an "inside view" and focusing on the constituents of the specific planned action rather than on the outcomes of similar actions already completed [39].

Table 1.Selected study related to construction project duration

Author(s)	Year	Country	Project Type	Approach
Bromilow	1969	Australia	Various buildings	BTC model- Simple linear Regression (SLR)
Bromilow	1974	Australia	Various buildings	BTC model- Simple linear Regression (SLR)
Arditi et al.	1985	Turkey	-	Survey
Walker	1994	Australia	Various buildings	Construction time performance (CTP)
Ng et al.	2001	Australia	Residential, Industrial, Educational, Recreational, Other	BTC model- Simple linear Regression (SLR)
Love et al.	2005	Australia	New build, refurbishment/renovation, fit out, new build/refurbishment	Weighted least squares (WLS), Survey
Ogunsemi and Jagboro	2006	Nigeria	Various buildings	BTC model- Simple linear Regression (SLR)
Faridi and El-Sayegh	2006	UAE	-	Survey
Stoy et al.	2007	Germany	Various buildings	Log-log regression
Sambasivan and Soon	2007	Malaysia	-	Survey
Sweis et al.	2008	Jordan	Residential buildings	Drewin's open conversion system
Abd El-Razek et al.	2008	Egypt	-	Survey
Fugar and Agyakwah-Baah	2010	Ghana	-	Survey
Dursun and Stoy	2012	Germany	Various buildings	Multiple linear regression (MLR)
Erdis	2013	Turkey	Public buildings	Decision trees (DTs), Artificial neural network (ANN), Support vector machines (SVM)
Guerrero et al.	2014	Spain	Various buildings	Multiple linear regression (MLR)
Gab-Allah et. al.	2015	Egypt	Various buildings	Artificial neural network (ANN)
Leu and Liu	2016	Taiwan	Industrial buildings	Artificial neural network (ANN)
Jarkas	2016	State of Kuwait	Office buildings	BTC Model- Simple linear regression (SLR), Multiple linear regression (MLR)
Vahdani et al.	2016	Iran	Construction	Neuro-fuzzy systems (NFS), Artificial neural network (ANN)
Jin et al.	2016	Korea	Multi housing buildings	Case-based reasoning (CBR)
Moret and Einstein	2016	Portugal	Tunnels, viaducts, cuts and embankments in rail lines	Decision aids for tunneling (DAT)
Flyvbjerg et al.	2016	Hong Kong	Roadworks	RCF
Batselier and Vanhoucke	2016	Belgium	Construction, building construction, commercial building construction	RCF, Monte Carlo simulation (MCS), Earned value management (EVM)

In response to this Kahneman and Frederick (2002) proposed a cure to the problem by namely taking an "outside view" on planned actions using distributional information from previous similar ventures [40].

RCF application basically requires the following five steps [41];

1) *Gathering projects planned and actual project data*: RCF application requires not only anticipated but also final values of the considered parameter. Otherwise a benchmarking would be impossible. The initial step of this study was to collect contract duration and actual duration data as well as the total construction area of completed public building projects by the Ministry of Environment and Urbanism in Turkey as the client. In total data from 643 building projects were collected. The projects were procured in accordance with the Public Procurement Law, nr. 4734 and completed in the period between 2003 and 2011 in various geographic regions in Turkey. The geographical areas include Adana, Ankara, Bursa, Gaziantep, Izmir, Malatya, Samsun and Trabzon. As a selection criterion, the information in the database relates only to duration of construction works and excluding repair, restoration etc.

2) *Identification of reference class(es) from past similar projects*: The phenomenon of 'optimism bias' describes the tendency of individuals to expect better than average outcomes from their actions [42]. Optimism bias 'uplift' expresses the level of the mentioned 'fallacy' in terms of percentage. A reference class represents similar types of projects. The number of projects in a reference class must be 'broad enough' to be statistically meaningful and cannot be too narrow; otherwise it would be difficult to establish valid optimism bias 'uplift'as each category is too small.Similarly reference classes cannot be too wide, because then some projects within each reference class are not comparable [33]. The key is to make sure, using statistical tests, benchmarking, and other analyses, that the overruns of projects within each reference class projects are statistically similar [41]. Batselier and Vanhoucke (2016) for instance identified three reference classes depending on the database as; construction, building construction and commercial building construction in order of increasing specificity and similarity [12]. Flyvbjerg (2004), Eythorsdottir (2012) also indicated that project type is the most significant selection criterion [33,37]. In this paper, the projects are split into different classes because of the large sample and the wide differences of duration as well as the variety of the intended use. For this purpose ranges for the 'project type' are considered as the most significant selection criterion for the various reference classes.

3) *Establishing a probability distribution for each reference class based on the gathered data*: In order to find the probability distribution of duration overrun for each reference class, the probability distribution for overrun/underrun has to be established. To ensure comparability, it is significant that the definition of estimated and actual is identical for all projects [33]. Duration overruns (%) were determined according to Equation (1),

$$I = \frac{(D_a - D_c)}{D_c} x100 \quad (1)$$

where *I = Duration overrun as percentage of contract duration,* D_a *= Actual duration of a project and* D_c *= Contract duration of a project.*

4) *Determining the required optimism bias uplift/reduction curve for each reference class*: It is possible to determine the required optimism bias "uplift" afterwards a probability distribution has been established for each considered reference class [33]. Required uplifts are established as a function of the level of risk one is willing to take as lower level of acceptable risk results in a higher required uplift. That is to say if in a reference class the average duration overrun is 10%, then to have a 50% chance of the actual duration being under the estimated duration and a 50% chance exceeding it, 10% uplift should be added to the estimate of the duration of a new project in the same reference class.

5) *Validating the obtained uplift/reduction values*: This step was unfortunately overlooked in the previous studies. Similar validation projects, which have not been considered in the established reference class(es), must be tested to see whether the obtained uplift/reduction values works.

RCF was first applied in practice in 2004 by Flyvbjerg on large British transportation projects [33]. Cost studies carried out by Flyvbjerg showed that many types of projects such as transportation projects, power plants, dams, water projects etc. follow a general pattern of under-estimation and overrun [43]. The method showed, for example, that with a willingness to accept 50% risk for cost overrun, the required uplift to produce realistic project cost forecasts for road projects and fixed links was 15% and 23% respectively. In both cases the lower the acceptable risk for overspend, the higher the uplift [33].

Consequently, RCF is a method that requires only one determinant as a basis to establish the uplift curve used for forecasting. As such it is very practical and useful compared to the other methods described earlier such as ANN, neuro-fuzzy etc. However despite the purported advantage of the method none of the previous published work on the implementation of RCF has included validation of the effectiveness of its produced forecasts. In other words, performed studies have determined uplift values but did not show whether these uplift values have been applied to produce realistic forecasts for new projects. In this study such validation is performed to bridge this knowledge gap.

4. Results

The characteristics of the collected projects is shown in Table 2. It shows that, half of the buildings were constructed in Izmir, Aegean region, and in Gaziantep, Southeastern Anatolia region. The procurement process is the same (e.g. open tender procedure and turnkey) for all provincial directorates. Approximately one of the three building projects total construction area is between 500 and 1,000 square meters. 34% of the projects were completed in time but 49% of the projects suffered from delays and 30% of the projects' duration overrun was more than 20%, which is quite considerable. Additional statistical analyses of the database are shown in Table 3.

Table 3 shows that the highest duration overrun is observed as 497.33% while the lowest duration underrun is-48.89%. The arithmetic mean of duration overrun of 24.18% is significant. Standard deviation value shows that the spread of values around the arithmetic mean is significantly wide and irregular. That is also clear from the high percentile for the coefficient of variation. The positive skewness value indicates that the tail of the distribution is extended to the right. The positive kurtosis value of 19.66 also indicates a highly-peaked distribution.

Table 2. General project characteristics, N=643

Category	Classification	Nr. of Projects	% of Projects
Provincial Directorate	Adana	61	9
	Ankara	65	10
	Bursa	56	9
	Gaziantep	143	23
	Izmir	183	28
	Malatya	69	11
	Samsun and Trabzon	66	10
Total Construction Area (m^2)	<500	90	14
	500-1,000	192	30
	1,000-2,000	119	18
	2,000-4,000	153	24
	> 4,000	89	14
Duration Overrun	> 20%	194	30
	10% - 20%	44	7
	0% - 10%	75	12
	= 0%	220	34
	< 0%	110	17

Table 3.Key statistical analysis information of the data, N=643

Parameter	Duration Overrun (%)
Lowest Value	-48.89
Highest Value	497.33
Arithmetic Mean	24.18
Standard Deviation	54.26
Coefficient of Variation (%)	224.43
Skewness	3.61
Kurtosis	19.66

The projects were grouped in reference classes in a following step. The educational buildings consist of kindergarten, elementary, secondary and high schools; the health service buildings consist of district hospitals, community and dental clinics and emergency services. The governmental buildings consist of the government offices in towns. The security buildings consist of the gendarmerie regional command buildings and guardhouse.

Table 4.Established reference classes, N=643

Category	RC Number	Classification	Nr. of Projects	% of Projects
Project Type	RC-1	Educational	515	80
	RC-2	Health service	77	12
	RC-3	Government	27	4
	RC-4	Security	24	4
		TOTAL	**643**	**100**

In an attempt to test the reliability of RCF to forecast project duration realistically, 95% of the data (data from 611 projects) were used to determine the probability distribution of the differences between contract duration and actual duration and accordingly to establish the optimism bias curves representing the project types. The data from the rest of the projects (32 projects) are then used to validate the results. Details of the modeling and test data are shown in Table 5.

The average duration overrun of the projects was found to be 25% for the modeling data and 16% for the test data.Figure 1 shows the duration overruns/underruns (%) of the modeling projects based on estimates.

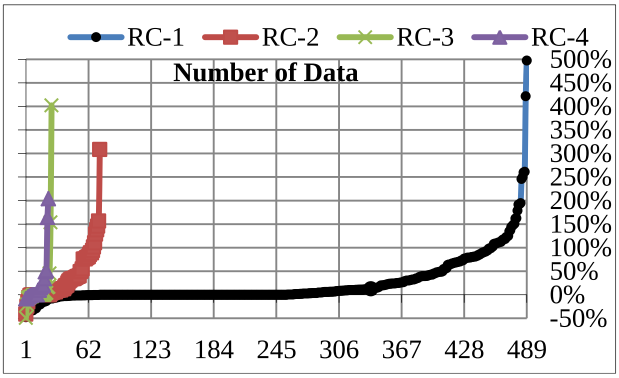

Figure 1. Distribution of duration overruns for the reference classes

For all duration overruns, the number of projects with a given maximum duration overrun was determined. Consequently, the probability distributions for the established reference classes were plotted as shown in Figure 2, which shows the distribution of duration overrun based on contract estimates for each reference class.For instance, 15% of the projects of RC-1 have a maximum duration overrun of 0% (meaning that actual durations were either equal or lower than the contract estimates) and that 71% of projects have a maximum overrun of 20%. For RC-2, 8% of the projects have a maximum overrun of 0% while 55% have a maximum overrun of 20%. For RC-3, 31% and 88% of the projects have a maximum overrun of 0% and 20% respectively. For RC-4, 30% and 74% of the projects have a maximum overrun of 0% and 20% respectively.

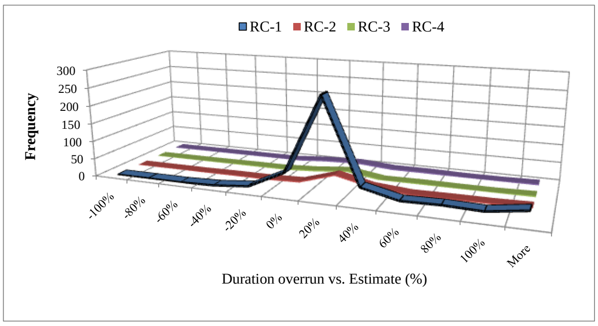

Figure2. Probability distribution of duration overruns for the reference classes

Table 5.Established reference classes, N=643

Ref. Class Nr.	Nr. of Modeling Data	Nr. of Modeling Data/Nr. of Total Data (%)	Average Duration Overrun (%)	Nr. of Test Data	Nr. of Test Data/Nr. of Total Data (%)	Average Duration Overrun (%)
RC-1	489	95	23	26	5	10
RC-2	73	95	35	4	5	30
RC-3	26	95	21	1	5	82
RC-4	23	95	25	1	5	30
TOTAL	**611**	**95**	**25**	**32**	**5**	**16**

Depending on the probability distribution of the reference classes shown in Figure 2, required uplifts for each reference class were determined. Acceptable chance of duration overrun between 0% and 100% for the collected data were initially identified. The required uplift value for each percentage of the identified range was then calculated. Finally the uplift values of the range between 10% and 50% were plotted as shown in Figure 3. Planners and promoters can choose the uplift that corresponds with the level of risk of duration overrun that they are willing to accept.

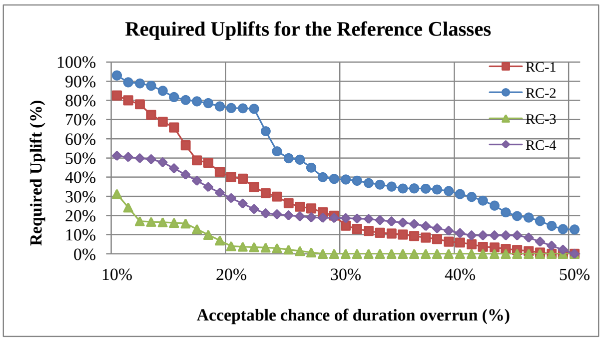

Figure 3. Required uplift as function of the maximum acceptable chance of duration overrun

Figure 3 shows that, for all reference classes except RC-2 (Health service), if decided that the acceptable risk of duration overrun for a new public building project should not be more than 50% (i.e. having up to 50% chance of being as the estimated duration), no uplift will be required. If the required acceptable risk of duration overrun is not more than 10% (i.e. having up to 90% chance of being as the estimated duration) then an uplift of 83%, 93%, 31% or 51% should be added to the project duration estimates in RC-1 to RC-4 respectively. Table 6 summarizes the exact required optimism bias uplifts for selected risk percentiles for the public building's reference classes based on Fig. 3.

Table 6. Required uplifts for the selected risk percentiles

Reference Class Nr.	Applicable optimism bias uplifts				
	10%	20%	30%	40%	50%
RC-1	83%	40%	15%	6%	0%
RC-2	93%	76%	39%	31%	13%
RC-3	31%	4%	0%	0%	0%
RC-4	51%	29%	19%	11%	0%

One shortcoming of the previous studies is that the application ended as soon as the uplift values were determined and they did not show whether these uplift values have been applied to produce realistic forecasts for new projects.The following section provides a validaton process.

5. Validation of RCF Forecasts

Data representing 5% of the projects in each reference class, randomly selected, were used as test data. In total they represent 32 projects with known actual duration as indicated in Table 5; 26 in RC-1, four in RC-2, one in RC-3 and one in RC-4. Test data are used for validating the method by adding the recommended uplift percentile for specific required risk levels for each of these projects based on their reference class and to compare the forecasts produced with the actual durations that are already known. Original duration overrun as well as the re-calculated duration overruns obtained by applying various uplift values on contract durations were calculated as percentages using Equation (1) for each test project.

Table 7.Original and uplift-based duration overruns for the test projects, N=32

No	Ref. Class Nr.	Original Duration Overrun (%)	Best Option for Re-Calculated Duration Overrun (%)	Acceptable Risk Providing Best Option (%)
1	RC-1	-39.64	-39.64	50
2		-35.16	-35.16	50
3		-31.11	-31.11	50
4		-25.71	-25.71	50
5		-18.71	-18.71	50
6		-15.00	-15.00	50
7		-9.00	-9.00	50
8		-8.00	-8.00	50
9		-6.67	-6.67	50
10		-5.74	-5.74	50
11		-3.67	-3.67	50
12		-1.67	-1.67	50
13		-0.63	-0.63	50
14		0.00	0.00	50
15		0.00	0.00	50
16		0.00	0.00	50
17		0.00	0.00	50
18		0.00	0.00	50
19		0.68	0.68	50
20		7.02	0.97	40
21		16.57	1.37	30
22		21.48	5.63	30
23		32.38	-5.44	20
24		51.90	8.50	20
25		103.13	11.00	10
26		233.18	82.06	10
27	RC-2	-34.45	-41.99	50
28		0.00	-11.50	50
29		35.54	-2.49	30
30		120.63	14.31	10
31	RC-3	81.67	38.68	10
32	RC-4	30.20	0.93	20

The risk level having the best option, on-time (closest to 0% duration overrun) was determined for each test project. These acceptable risk distributions of the duration overruns of the test projects are shown in Table 7 and the results obtained are evaluated in Fig. 4 which shows that the average duration overruns based on the planned duration without applying uplifts are 10.22%, 30.43%, 81.67% and 30.20% whilst the average duration overruns after applying uplifts on the contract duration are -3.69%, -10.42%, 38.68% and 0.93% from RC-1 to RC-4 respectively.

These values indicate that the forecast accuracy of the average overrun has improved (or become more realistic) in all reference classes from a range of 38.13% to 6.37%. Table 7 shows that in the case of

RC-1, the accurate forecasts of the actual final duration produced using RCF for 19 out of the 26 test projects (i.e. 73%), are achieved using 50% acceptable risk of overrun. Table 7 also shows that in the case of RC-2, accurate forecasts for two out of the four projects were produced using up to 30% acceptable risk. Similar accurate forecasts are also produced for projects in the reference classes RC-3 and RC-4. The average acceptable risks for the best options obtained based on RC-1 to RC-4 are 42.69%, 35.00%, 10.00% and 20.00% respectively. Consequently, an evaluation of the 32 test projects in total indicates that the original average duration overrun of 15.60% have been obtained as -3.06% using the required uplift values.

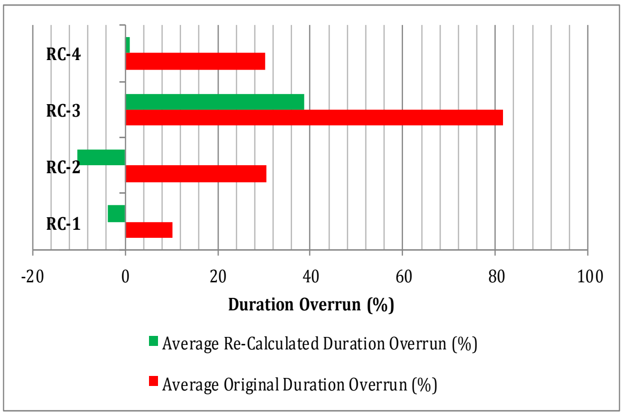

Figure 4. Evaluation of the results obtained from test projects

The average acceptable risk of the 32 test projects is also obtained as 40.00%. The test results indicate that for the majority of building projects, durations are underestimated, which means that planners were optimistic and the requirement based on realistic duration indicates that some uplifts on the original estimated durations are required.

6. Discussion

The study was limited to analyzing only the relationship between contractual and actual construction time using public building projects in Turkey as a case. Contract and actual durations under 100 days were considered to have limited scope and complexity.

The analysis has shown that almost half of the projects used in this study had suffered delays which in a way question the effectiveness of the procedure and guidelines currently used based on the Public Procurement Law, nr. 4734. The distribution of the duration overrun values based on RC-1, school projects, which consist of 80% of the total modeling data, is the widest among the established reference classes with a range between-47.64% and 497.33% and has an average duration overrun of 23%. However among the reference classes, the highest average duration overrun of 35% is found in RC-2, hospital projects. This may seem surprising and not expected but planners are usually overconfident when they estimate project duration and even more so when it comes to the health service and hence tend to produce optimistic and unrealistic estimates. The results seem to be in agreement with the underlying assumption of the theory on which RCF is based. The results also show that the narrowest range of duration overrun values is between -8.73% and 204.01% in RC-4, security projects, having 25% average duration overrun. The least average duration overrun is obtained as 21% from RC-3, government projects.

The required optimism bias uplifts for the established reference classes indicate for example that when the accepted risk of duration overrun for a new public building project is not more than 50% (i.e. having up to 50% chance to be exactly as the estimated duration), none of the estimates in any of the reference classes, except RC-2 will require an uplift. On the other hand if more certainty is desired and the accepted risk of the actual duration exceeding the estimate is not more than 30% (i.e. having up to 70% chance for the estimate to be accurate), it would be necessary to use an uplift of 39% on the duration estimate of projects in RC-2, 19% and 15% in RC-4 and RC-1, and 0% in RC-3. Furthermore for lower acceptable levels of risks, uplift values decrease in RC-2, RC-1, RC-4 and RC-3 respectively. For example, if it was decided that the acceptable risk of duration overrun should not be more than 10%, uplifts of 93%, 83%, 51% or 31% should be used on estimates in RC-2 to RC-3 respectively. This means that different project types cause significant differences based on duration estimates. It is clear that the most risky project type is health service while governmental projects feature lowest risk. This is unexpected because as the total construction area increases, duration overrun is expected to increase due to the complexity in work and size of the project. However, based on the aforementioned risk ranking, average total construction area is; 2102, 2076, 2982 and 1907 m^2 from RC-2 to RC-3 respectively. This means that the governmental projects, featuring lowest risk have the highest total construction area in average. One explanation for this is that planners may underestimate the cost and duration of small buildings as being simple and straightforward. In general however the study has shown that positive bias uplift values are required for all acceptable levels of risk that is equal or lower than 50% in all reference classes of buildings. This seems to justify the assumption that planners are usually optimistic when estimating contract duration.

Moreover the test results indicate that the acceptable risk of duration overrun of the projects in RC-1, RC-2, RC-4 to RC-3 decreases respectively. For instance, the average acceptable risk value, obtained from the best options of the duration overruns for the test projects is calculated up to be 43%, 35%, 20% and 10% from

RC-1 to RC-3 respectively. These results indicate that the risk margin in school and health service projects is higher than those produced for security and government projects.

As previously mentioned, the studies related RCF has so far been generally focused on cost forecasting. Among the few studies related RCF to forecast project duration were discussed below;

- Moret and Einstein (2016) established different RC's including tunnels, viaducts, cuts & embankments with a total of 55 projects from Portugese rail lines [10]. Project durations in a wide range from 1-day to 365-days were considered. However the durations under 100 days were considered to have limited scope and complexity in the current study. Moret and Einstein (2016) observed the largest increases in duration of the cuts & embankments (94%). However it cannot be expressed that at which level 94% increase was obtained. The current study indicates for example that if it was decided that the acceptable risk of duration overrun should not be more than 10%, uplift of 93% should be used on estimates in health service buildings.

- Database of the study performed by Batselier and Vanhoucke (2016) consisted of only 56 Belgian projects, from many different companies from various sectors conversely [12]. Three reference classes were established as; construction, building construction and commercial building construction. This approach specificly cannot reflect a different perspective since it respresents almost all the construction types. The current study on the other hand is only focused on the public buildings and classified them. Although Batselier and Vanhoucke (2016) mentioned that the RCF indeed performs best for both cost and time forecasting, their database is dispersed and contradicts with the 2nd step described at the Methodology section.

Therefore the main contribution of this paper is limiting project duration in a logical range and comparing specific and certain project types.Consequently, it has been proven that RCF produces realistic results by producing forecast with an average of 120% from the actual forecasts based on the validation projects and that by accounting of the uncertainties using various uplift values.

7. Conclusion

So far RCF applications have been generally used in cost forecasting of infrastructure projects. This paper describes the pioneering application of RCF method for forecasting duration of public building projects. As such the paper provides useful scientific contribution to knowledge on the application of RCF and to construction planning practice by investigating the accuracy of its forecasts of duration in building projects.

The current study analysed the project duration data of a total of 643 completed public building projects carried out in Turkey. The collected data consisted of actual duration and contract duration as well as total construction area of each project. Four different reference classes; RC-1 to RC-4 representing 'project type' were used. Data from 611 projects were used to establish the optimism bias uplift values for the reference classes and data from the remaining 32 projects were used as control projects to test the reliability of the method.

RCF forecasts produced based on the established optimism bias uplift curves of the reference classes indicated that for acceptable risk of duration overrun of 50% or more, none of the reference classes except RC-2, health service, required an uplift, which means that the original estimates will be adequate enough to produce durations similar to the actual duration. However for acceptable risk of duration overrun of not more than 30%, original estimates of duration of projects would require uplift values of 15%, 19%, 0% and 39% in RC-1, RC-4, RC-3 and RC-2 respectively. It was also mentioned that for the lower levels of acceptable risk of up to 10%, uplift required for project duration estimates in RC-3 to RC-2 increases from 31% to 93% respectively. This indicates that the health service buildings are exposed to highest risks while the government buildings are exposed to the lowest. The average construction areas of the buildings in these reference classes have also been investigated and unexpectedly, highest average construction area belongs to the low-risk government buildings. The application of RCF on a big sample of public project buildings in Turkey has shown that reasonable and realistic prediction of final project duration can be achieved for various levels of acceptable risk by adding the appropriate uplift to the original duration estimate produced by planner.

Duration forecasting is as important as cost forecasting for the construction projects. This study has shown that the RCF method, which is associated with cost forecasting, can also be used in forecasting the project duration and that it produces realistic results.

Acknowledgment

The authors would like to thank "The Scientific and Technical Research Council of Turkey (TUBITAK)" for the contribution to the corresponding author with "2219-International Postdoctoral Research Scholarship Programme".

References

[1] Laptalı, E., Bouchlaghem, N. M., Wild, S. 1996. An Integrated Computer Model of Time and Cost Optimisation. 12th Annual ARCOM Conference, 11-13 September, Sheffield Hallam University, UK, 133-139.

[2] Owolabi, J.D., Amusan, L.M., Oloke, C.O., Olusanya, O., Tunji- Olayeni, P., Owolabi, D., Peter, J., Omuh, I. 2014. Causes and Effect of Delay on Project Construction Delivery Time. International Journal of Education and Research, 2(4), 197-208.

[3] Faridi, A.S., El-Sayegh, S.M. 2006. Significant Factors Causing Delay in the UAE Construction Industry. Construction Management and Economics, 24(11), 1167-1176.

[4] Sambasivan, M., Soon, Y.W. 2007. Causes and Effects of Delays in Malaysian Construction Industry. International Journal of Project Management, 25(5), 517-526.

[5] Fugar, F.D.K., Agyakwah-Baah, A.B. 2010. Delays in Building Construction Projects in Ghana. Australasian Journal of Construction Economics and Building, 10(1/2), 103-116.

[6] Couto, J.P., Teixeira, J.C. 2007. The Evaluation of the Delays in the Portuguese Construction. CIB World Building Congress `Construction for Development`, 14-18 May, Cape Town, South Africa, 292-301.

[7] Ng, T., Mak, M.M.Y., Skitmore, M., Lam, K.C., Varnam, M. 2001. The Predictive Ability of Bromilow's Time–Cost Model. Construction Management and Economics, 19(2), 165-173.

[8] Zujo, V., Diana, C.P., Vejzovic, A.B. 2010. Contracted Price Overrun as Contracted Construction Time Overrun Function. Technical Gazette, 17(1), 23-29.

[9] Construction Sector Report. 2016. The Turkish Employers Association of Construction Industries (INTES).

http://intes.org.tr/ti/779/0/Insaat-Sektor-Raporu-(Temmuz---2016).php

(Access date: 20.09.2017).

[10] Moret, Y., Einstein, H.H. 2016. Construction Cost and Duration Uncertainty Model: Application to High-Speed Rail Line Project. Journal of Construction Engineering and Management, ASCE, 142(10), 05016010.

[11] Flyvbjerg, B., Hon, C.K., Fok, W.H. 2016. Reference Class Forecasting for Hong Kong's Major Roadworks Projects. Proceedings of the Institution of Civil Engineers - Civil Engineering, 169(6), 17-24.

[12] Batselier, J., Vanhoucke, M. 2016. Practical Application and Empirical Evaluation of Reference Class Forecasting for Project Management. Project Management Journal, 47(5), 36-51.

[13] Arditi, D., Akan, G.T., Gurdamar, S. 1985. Reasons for Delays in Public Projects in Turkey. Construction Management and Economics, 3(2), 171-181.

[14] Erdis, E. 2013. The Effect of Current Public Procurement Law on Duration and Cost of Construction Projects in Turkey. Journal of Civil Engineering and Management, 19(1), 121–135.

[15] Sweis, G., Sweis, R., Abu Hammad, A., Shboul, A. 2008. Delays in Construction Projects: The Case of Jordan. International Journal of Project Management, 26(6), 665-674.

[16] Abd El-Razek, M.E., Bassioni, H.A, Mobarak, A.M. 2008. Causes of Delays in Building Construction Projects in Egypt. Journal of Construction Engineering and Management, ASCE, 134(11), 31-841.

[17] Vandevoorde, S., Vanhoucke, M. 2006. A Comparison of Different Project Duration Forecasting Methods Using Earned Value Metrics. International Journal of Project Management, 24(4), 289-302.

[18] Bromilow, F.J. 1969. Contract Time Performance: Expectations and the Reality. Building Forum, 1(3), 70-80.

[19] Bromilow, F.J. 1974. Measurement and Scheduling of Construction Time and Cost Performance in the Building Industry. The Chartered Builder, 10(9), 79-82.

[20] Odabasi, E. 2009. Models for estimating construction duration: An application for selected buildings on the METU campus. MSc thesis, Middle East Technical University, Ankara, Turkey.

[21] Dursun, O., Stoy, C. 2011. Time-Cost Relationship of Building Projects: Statistical Adequacy of Categorization with Respect to Project Location. Construction Management and Economics, 29(1), 97-106.

[22] Ogunsemi, D.R., Jagboro, G.O. 2006. Time-Cost Model for Building Projects in Nigeria. Construction Management and Economics, 24(3), 253–258.

[23] Walker, D.H.T. 1994. An investigation into the factors that determine building construction time performance. PhD dissertation, Royal Melbourne Institute of Technology, Melbourne, Australia.

[24] Love, P.E.D., Tse, R.Y.C., Edwards, D.J. 2005. Time-Cost Relationships in Australian Building Construction Projects. Journal of Construction Engineering and Management, ASCE, 131(2), 187-194.

[25] Stoy, C., Pollalis, S., Schalcher, H. 2007. Early Estimation of Building Construction Speed in Germany. International Journal of Project Management, 25(3), 283–289.

[26] Dursun, O., Stoy, C. 2012. Determinants of Construction Duration for Building Projects in Germany. Engineering, Construction and Architectural Management, 19(4), 444-468.

[27] Guerrero, M.A., Villacampa, Y., Montoyo, A. 2014. Modeling Construction Time in Spanish Building Projects. International Journal of Project Management, 32(5), 861–873.

[28] Gab-Allah, A.A., Ibrahim, A.H., Hagras, O.A. 2015. Predicting the Construction Duration of Building Projects Using Artificial Neural Networks. International Journal of Applied Management Science, 7(2), 123-141.

[29] Leu, S.S., Liu, C.M. 2016. Using Principal Component Analysis with a Back-Propagation Neural Network to Predict Industrial Building Construction Duration. Journal of Marine Science and Technology-Taiwan, 24(2), 82-90.

[30] Jarkas, A.M. 2016. Predicting Contract Duration for Building Construction: Is Bromilow's Time-Cost Model a Panacea? Journal of Management in Engineering, ASCE, 32(1), 05015004.

[31] Vahdani, B., Mousavib, S.M., Mousakhanic, M., Hashemid, H. 2016. Time Prediction Using a Neuro-Fuzzy Model for Projects in the Construction Industry. Journal of Optimization in Industrial Engineering, 9(19), 97-103.

[32] Jin, R.Z., Han, S., Hyun, C.T., Cha, Y. 2016. Application of Case-Based Reasoning for Estimating Preliminary Duration of Building Projects. Journal of Construction Engineering and Management, ASCE, 142(2), 04015082.

[33] Flyvbjerg, B. 2004. Procedures for Dealing with Optimism Bias in Transport Planning, Guidance Document in association with COWI. Report nr. 58924, The British Department for Transport, UK.

[34] Flyvbjerg, B., Holm, M.K.S., Buhl, S.L. 2005. How (In)accurate are Demand Forecasts in Public Works Projects? The Case of Transportation. Journal of the American Planning Association, 71(2), 131-146.

[35] Flyvbjerg, B. 2007. Eliminating Bias Through Reference Class Forecasting and Good Governance. Report nr. 17, Norwegian University of Science and Technology (NTNU), Trondheim, Norway.

[36] Salling, K.B., Leleur, S. 2012. Modelling of Transport Project Uncertainties: Feasibility Risk Assessment and Scenario Analysis. European Journal of Transport and Infrastructure Research (EJTIR), 12(1), 21-38.

[37] Eythorsdottir, E.O. 2012. Reference class forecasting method used in Icelandic transportation infrastructure projects. MSc thesis, Reykjavík University, Reykjavík, Iceland.

[38] Kahneman, D. 2002. Maps of Bounded Rationality: A perspective on intuitive judgment and choice. Prize Lecture, The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel, Princeton University, USA.

[39] Kahneman, D. 1994. New Challenges to the Rationality Assumption. Journal of Institutional and Theoretical Economics, 150(1), 18-36.

[40] Kahneman, D., Frederick, S. 2002. Representativeness Revisited: Attribute Substitution in Intuitive Judgment, in T. Gilovich, D. Griffin, and D. Kahneman, 2002 eds., Heuristics and biases: The Psychology of Intuitive Judgment, Cambridge University Press, NY, US, 49–81.

[41] Flyvbjerg, B. 2006. From Nobel Prize to Project Management: Getting Risks Right. Project Management Journal, 37(3), 5-15.

[42] De Reyck, B., Grushka-Cockayne, Y., Fragkos, I., Harrison, J., Read, D. 2015. Optimism Bias Study Recommended Adjustments to Optimism Bias Uplifts: Final Report. Department for Transport, UK.

[43] Flyvbjerg, B. 2005. Design by Deception: The Politics of Megaproject Approval. Harvard Design Magazine, 50-59.