UCTEA Turkish Chamber of Civil Engineers

Teknik Dergi

Technical Journal

Volume 31 Issue 2 March 2020

TEKNİK DERGİ PUBLICATION PRINCIPLES

Teknik Dergi is a scientific and technical journal indexed by the Science Citation Index Expanded. Annually six issues are published, three in Turkish in the months of January, May and September, three in English in March, July and November. Its main principles of publication are summarized below:

1. Articles reporting original scientific research and those reflecting interesting engineering applications are accepted for publication. To be classified as original, the work should either produce new scientific knowledge or add a genuinely new dimension to the existing knowledge or develop a totally new method or substantially improve an existing method.
2. Articles reporting preliminary results of scientific studies and those which do not qualify as full articles but provide useful information for the reader can be considered for publication as technical notes.
3. Discussions received from the readers of the published articles within three months from publication are reviewed by the Editorial Board and then published together with the closing remarks of the author.
4. Manuscripts submitted for publication are evaluated by two or three reviewers unknown to the authors. In the light of their reports, final decision to accept or decline is taken by the Editorial Board. General policy of the Board is to get the insufficient manuscripts improved in line with the reviewers' proposals. Articles that fail to reach the desired level are declined. Reasons behind decisions are not declared.
5. A signed statement is taken from the authors, declaring that the article has not been published as a "journal article or book chapter". In case the Editorial Board is in the opinion that the article has already been published elsewhere with minor changes or suspects plagiarism or a similar violation of ethics, then not only that article, but none of the articles of the same authors are published.
6. Papers reporting works presented as conference papers and developed further may be considered for publication. The conference it was presented to is given as a footnote in the first page.
7. Additionally, a document signed by all authors, transferring the copyright to UCTEA Chamber of Civil Engineers is submitted together with the manuscript.

UCTEA Turkish Chamber of Civil Engineers

Teknik Dergi

Technical Journal

Volume 31 Issue 2 March 2020

UCTEA ***(TMMOB)***
Turkish Chamber of Civil Engineers ***(İnşaat Mühendisleri Odası)***
Necatibey St. No: 57, Kızılay 06440 Ankara, Turkey
Tel: +90.312.294 30 00 - Faks: +90.312.294 30 88
E-mail: imo@imo.org.tr - www.imo.org.tr

Publisher ***(Sahibi)*****:**
Cemal GÖKÇE
On behalf of UCTEA Turkish Chamber of Civil Engineers

Administrative Officer ***(Yazı İşleri Müdürü)*****:**
Bahaettin SARI

Volume 31 - Issue 2 - March 2020 *(Cilt 31 - Sayı 2 - Mart 2020)*

Published bi-monthly. Local periodical. *(İki ayda bir yayınlanır, yerel süreli yayın)*

Date of Print: 01 March 2020 *(Baskı Tarihi: 1 Mart 2020)*

Number of copies: 1.000 *(1.000 adet basılmıştır)*

ISSN: 1300-3453

Printed by ***(Baskı)*****:**
Yorum Basın Yayın Sanati Ltd. Şti.
Başkent Organize Sanayi Bölgesi Recep Tayyip Erdoğan Bulvarı No: 12
Malıköy - Sincan /Ankara - Tel: 0.312.395 21 12

UCTEA Turkish Chamber of Civil Engineers

Teknik Dergi

Teknik Dergi is indexed by

- Science Citation Index Expanded
- Scopus
- Journal Citation Reports / Science Edition
- Engineering Index
- Concrete Abstracts (American Concrete Institute)
- National Technical Information Service (US NTIS)
- CITIS
- Ulrich’s International Periodical’s Directory
- TÜBİTAK / ULAKBİM

Teknik Dergi is a peer reviewed open access periodical publishing papers of original research and interesting practice cases. It addresses both the research community and the practicing engineers.

Reviewers:

This list is renewed each year and includes reviewers who served in the last two years of publication.

Ayda Şafak AĞAR ÖZBEK
Perviz AHMEDZADE
Ragıp AKBAŞ
Sami Oğuzhan AKBAŞ
Rıfat AKBIYIKLI
Özge AKBOĞA KALE
Burcu AKÇAY ALDANMAZ
Cihan Taylan AKDAĞ
Cem AKGÜNER
Adem AKPINAR
Muhammet Vefa AKPINAR
Atakan AKSOY
Hafzullah AKSOY
Güzin AKYILDIZ ALÇURA
Zuhal AKYÜREK
Fatih ALEMDAR
Pelin ALPKÖKİN
Sinan ALTIN
Selim ALTUN
Adlen ALTUNBAŞ
Ahmet Can ALTUNIŞIK
Egemen ARAS
Fuat ARAS
Davit ARDITI
Ergin ARIOĞLU
Deniz ARTAN İLTER
Ali Osman ATAHAN
Hakan Nuri ATAHAN
Shady ATTIA
Abdullah AVEY
İsmail AYDIN
Mustafa Tamer AYVAZ
Ela BABALIK
Can Elmar BALAS
Lale BALAS
Selim BARADAN
Türkay BARAN
Bekir Oğuz BARTIN
Cemal BAŞARAN
Zeynep BAŞARAN BUNDUR
Özgür BAŞKAN
Cüneyt BAYKAL
İdris BEDİRHANOĞLU
Mehmet BERİLGEN
Saadet Arzu BERİLGEN
Niyazi Özgür BEZGİN
Selçuk BİLDİK
Senem BİLİR MAHÇİÇEK
Barış BİNİCİ
Ahmet BİRİNCİ
İlknur BOZBEY
Zafer BOZKUŞ
Burcu BURAK BAKIR
Halil İbrahim BURGAN
Yusuf CALAYIR
Erdem CANBAY
Zekai CELEP
Cihan CENGİZ
Halim CEYLAN
Ömer CİVALEK
Mustafa CÖMERT
Ali Fırat ÇABALAR
Barlas Özden ÇAĞLAYAN
Özgür ÇAKIR
Melih ÇALAMAK
Gülben ÇALIŞ
Erkan ÇELEBİ
Kutay ÇELEBİOĞLU
Ahmet Ozan ÇELİK
Oğuz Cem ÇELİK
Osman Nuri ÇELİK
Semet ÇELİK
Hilmi Berk ÇELİKOĞLU
Kemal Önder ÇETİN
Mecit ÇETİN
Reha ÇETİNKAYA
Gökhan ÇEVİKBİLEN
Mesut ÇİMEN
Safiye Feyza ÇİNİCİOĞLU
Erdal ÇOKÇA
Şevket ÇOKGÖR
Atilla DAMCI
Yakup DARAMA
Kutlu DARILMAZ
Cem DEMİR
Uğur DEMİR
Ender DEMİREL
Mehmet Cüneyd DEMİREL
Fatih DİKBAŞ
Seyyit Ümit DİKMEN
İrem DİKMEN TOKER
Ali Ersin DİNÇER
Ahmet Anıl DİNDAR
Emrah DOĞAN
Nurhan ECEMİŞ ZEREN
Özgür EKİNCİOĞLU
Alper ELÇİ
Şebnem ELÇİ
Murat Altuğ ERBERİK
Saffet ERDOĞAN
Esin ERGEN PEHLEVAN
Ayşen ERGİN
Gökmen ERGÜN
Ebru ERİŞ
Esra Ece ESELLER BAYAT
Tuğba ESKİŞAR TEFÇİ
Burak FELEKOĞLU
Okan FISTIKOĞLU
Antonio FORMISANO
Nuray GEDİK
Abdullah GEDİKLİ
Ergun GEDİZLİOĞLU
Mohammad Ali GHORBANİ
Konuralp GİRGİN
Zehra Canan GİRGİN
Ilgın GÖKAŞAR
Çağlar GÖKSU
Burcu GÜLDÜR ERKAL
Fazlı Erol GÜLER
Hakan GÜLER
Ilgın GÜLER
Zeynep GÜLERCE
Taylan GÜNAY
Necmettin GÜNDÜZ
Abdurrahman GÜNER
Samet GÜNER
Ülker GÜNER BACANLI
Mehmet Şükrü GÜNEY
Tuba GÜRBÜZ BÜYÜKKAYIKÇI
Gürkan Emre GÜRCANLI
Aslı Pelin GÜRGÜN
İpek GÜRSEL DİNO
Gürşans GÜVEN IŞIN
Soner HALDENBİLEN
Murat HAMDERİ
Ufuk HANCILAR
Ingo A. HANSEN
Abdul HAYIR
Nejan HUVAJ SARIHAN
Metin HÜSEM
Zeynep IŞIK
Sabriye Banu İKİZLER
Eren İNCİ
Pınar İNCİ KOÇAK
Erdal İRTEM
Nihat KABAY
Sedat KABDAŞLI
Volkan KAHYA
Mehmet Rifat KAHYAOĞLU
Volkan KALPAKÇI
Alper KANYILMAZ
Murat KARACASU
Tanay KARADEMİR
Erhan KARAESMEN
Ali KARAİPEKLİ
Himmet KARAMAN
Mustafa KARAŞAHİN
Zülküf KAYA
İlker KAZAZ
Cevza Melek KAZEZYILMAZ ALHAN
Mustafa Kubilay KELEŞOĞLU
Elçin KENTEL
Mustafa Erol KESKİN
Havvanur KILIÇ
İsmail Emrah KILIÇ
Sami And KILIÇ
Fahriye KILINÇKALE
Ufuk KIRBAŞ
Veysel Şadan Özgür KIRCA
Gökhan KİRKİL
Niyazi Uğur KOÇKAL
Önder KOÇYİĞİT
Baha Vural KÖK
Mete KÖKEN
Fuat KÖKSAL
Ali Ümran KÖMÜŞÇÜ
Şerife Yurdagül KUMCU
Akif KUTLU
Semih KÜÇÜKARSLAN
Abdullah KÜRKÇÜ
Hilmi LUŞ
Kasım MERMERTAŞ
Mehmet Murat MONKUL
Yetiş Şazi MURAT
Elif OĞUZ
Didem OKTAY
Volkan OKUR
Mehmet Hakkı OMURTAG
Sema ONURLU
Engin ORAKDÖĞEN
Şeref ORUÇ
Okan ÖNAL
Akın ÖNALP
Halil ÖNDER
Aybike ÖNGEL
Bihrat ÖNÖZ
Ali Hakan ÖREN
Bergüzar ÖZBAHÇECİ
Ceyhun ÖZÇELİK
İlker ÖZDEMİR
Murat ÖZEN
Pelin ÖZENER
Abdullah Tolga ÖZER
Eren Arman ÖZGÜVEN
Hakkı Oral ÖZHAN
Yener ÖZKAN
M. Hulusi ÖZKUL
Zeynep Huri ÖZKUL BİRGÖREN
Beliz ÖZORHON ORAKÇAL
Sadık ÖZTOPRAK
Turan ÖZTURAN
Baki ÖZTÜRK
Hasan Tahsin ÖZTÜRK
Mustafa ÖZUYSAL
Tolga Yılmaz ÖZÜDOĞRU
Polat ÖZYİĞİT
Gülizar ÖZYURT TARAKÇIOĞLU
Nilüfer ÖZYURT ZİHNİOĞLU
Onur PEKCAN
Bekir Yılmaz PEKMEZCİ
Cengiz POLAT
Şamil Şeref POLAT
Gül POLAT TATAR
Selim PUL
Selçuk SAATÇİ
Selman SAĞLAM
Mehmet SALTAN
Altuğ SAYGILI
Neslihan SEÇKİN
Serdar SELAMET
Alper SEZER
Faiz Uddin Ahmed SHAİKH
Osman SİVRİKAYA
Serdar SOYÖZ
Aleksandar STEVANOVİC
Ayşe Filiz SUNAR
Erol ŞADOĞLU
Remzi ŞAHİN
Yuşa ŞAHİN
Mustafa ŞAHMARAN
Nermin ŞARLAK
Burak ŞENGÖZ
Aynur ŞENSOY ŞORMAN
Ali Arda ŞORMAN
Ali Ünal ŞORMAN
Özcan TAN
Ali Hamza TANRIKULU
Kürşat TANRIÖVEN
Serhan TANYEL
Taha TAŞKIRAN
Gökmen TAYFUR
İlker TEKİN
Beytullah TEMEL
Berrak TEYMUR
H. Onur TEZCAN
Mesut TIĞDEMİR
Şahnaz TİĞREK
Vedat TOĞAN
Nabi Kartal TOKER
Nuray TOKYAY
Ali TOPAL
Cem TOPKAYA
Selçuk TOPRAK
Ahmet TORTUM
Gökçe TÖNÜK
Nursu TUNALIOĞLU
Kağan TUNCAY
Eda TURAN
Gürsoy TURAN
Kaan TÜRKER
Cüneyt TÜZÜN
Eren UÇKAN
Latif Onur UĞUR
Mehmet Fevzi UĞURYOL
Berna UNUTMAZ
Volkan Emre UZ
Deniz ÜLGEN
Aslı ÜLKE KESKİN
Cüneyt VATANSEVER
Syed Tanvir WASTI
Nazmiye YAHNİOĞLU
Ahmet YAKUT
Cem YALÇIN
Aslı YALÇIN DAYIOOĞLU
İsmail Özgür YAMAN
A. Melih YANMAZ
Mert Yücel YARDIMCI
Ufuk YAZGAN
Anıl YAZICI
Halit YAZICI
Kasım YENİGÜN
İrem Zeynep YILDIRIM
Mehmet YILDIRIMOĞLU
Osman YILDIZ
Çetin YILMAZ
Fatih YILMAZ
Koray Kamil YILMAZ
M. Tuğrul YILMAZ
Mehmet YILMAZ
Murat YILMAZ
Veysel YILMAZ
Yüksel YILMAZ
Fatih YONAR
Recep YURTAL
İsmail YÜCEL
Ercan YÜKSEL
Yeliz YÜKSELEN AKSOY
Nabi YÜZER
Ahmet Şahin ZAİMOĞLU

UCTEA Turkish Chamber of Civil Engineers

Teknik Dergi

Volume: 31 Issue: 2 March 2020

CONTENTS

Teknik Dergi, 2020 9813-9832, Paper 568

Factors Affecting the Capacity Utilization of Road Freight Transport in Turkey

Murat ÖZEN[1]
Muhammad FAYYAZ[2]
Hediye TÜYDEŞ YAMAN[3]

ABSTRACT

Road freight accounts for 89% of the inland freight transportation in Turkey. This study examined the factors affecting capacity utilization in Turkey with a specific focus on load factors and empty running. These two decisions, the load factor and the choice of running empty, are modeled jointly in a Heckman model. The model results showed that rigid and articulated trucks represent relatively different transport markets. Their age and the distance profile of the shipments affect the load factor and the probability of running empty. These effects are further examined for different commodity types and by testing different econometric specifications.

Keywords: Capacity utilization, load factor, empty running, heckman model, commodity types.

1. INTRODUCTION

Road dominates the inland freight transportation in Turkey with 89% market share [1]. The situation is quite similar in many other countries [2]. For instance, in the European Union-28 region, 75% of the inland freight is transported by trucks [3]. Although there have been several policies to reduce share of road freight, the dominance of trucking is expected to continue due to its flexibility and door-to-door delivery [4]. This problem is more critical in developing countries, such as Turkey, in which other alternatives are not competitive due to different reasons (i.e. shorter hauling distances, lack of intermodal freight infrastructure and legislation, etc.). Therefore, it becomes more important to evaluate efficiency of road freight

Note:

- This paper has been received on April 27, 2018 and accepted for publication by the Editorial Board on February 04, 2019.
- Discussions on this paper will be accepted by May 31, 2020.

- https://dx.doi.org/10.18400/tekderg.413807

1 Mersin University, Department of Civil Engineering, Mersin, Turkey - ozen.murat@mersin.edu.tr - https://orcid.org/0000-0002-1745-7483

2 Institute of Transport and Logistics Studies, The University of Sydney, Sydney, Australia - muhammad.fayyaz@sydney.edu.au - https://orcid.org/0000-0003-1075-5688

3 Middle East Technical University, Department of Civil Engineering, Ankara, Turkey - htuydes@metu.edu.tr - https://orcid.org/0000-0003-2053-992X

movements to reach more efficient and environmentally friendly freight transportation system.

Capacity utilization (load factor) is a key factor in assessing efficiency of freight transportation. Load factor can be defined as ratio of actual weight of goods to the maximum weight that can be transported on a fully loaded trip [7]. It is a weight based measure and it may underestimate actual utilization of vehicle in some sectors, such as automobile, food and parcels where utilization of vehicles are limited by volume rather than weight [8]. Abate and Kveiborg [5] stated that, the capacity utilization of the trucks depends on several factors including characteristics of a trip (length, geographical area that it takes place, etc.), type of truck (operator type, size, etc.) as well as type of goods carried. The decision on how much to load or whether the truck should run empty represents an economic trade-off for the truck operator. Empty running occurs due to the unavailability of a load at the time of the return trip. However, it may also occur, if carriers have to deliver to several locations during a trip chain. Sometimes it is unavoidable, since freight trips are generally performed from point of production centers to consumption centers [5-6]. Despite an increase in revenue, a return load may also represent additional costs in terms of detouring and waiting costs, which may outweigh the additional revenue [9]. As a result, the frequency of empty runs is not random across the regions. For example, for importing regions it is harder to find matching return trips.

It is essential to represent this decision problem in a joint framework in which the joint decision of running empty and choosing a specific load-factor was modeled as a sample selection problem [10-12]. The approach was applied to Danish heavy truck trips, for which it was clearly shown that there are correlations between the two stochastic processes. The objective of this study is to examine the factors affecting capacity utilization in road freight transportation for Turkey by investigating the joint decision problem of running empty (a binary choice) and having a certain load factor. Using disaggregate data collected by the Turkish General Directorate of Highways (TGDH) from 2007 to 2011, this study has applied a joint estimation methodology for empty running and the load factor decision.

The main contribution of the study is the use of generally accepted models within a developing country context such that dynamics of empty runs in Turkey are determined. In a broader perspective, the paper contributes to the understanding of truck capacity utilization trade-offs in developing countries where there is a high share of road transport, an old fleet of trucks, lower GDP, and a different degree of urbanization compared to industrialized countries. Furthermore, the findings of this study have relevance for other developing countries, where there are few alternatives to road transport due to lack of rail and shipping infrastructure. A limitation of the study is the slightly short term of data availability, which was still long enough to derive statistically significant patterns, which can be used to develop more efficient trucking sector policies towards a low carbon development structure for Turkey.

The next section of this paper describes the data used in this study; which includes data characteristics and limitations. The section thereafter describes the econometric framework which consists of factors affecting load factor and empty running, and model specification based on the Heckman model. The following section presents the overall results for the various models. The concluding section provides some policy implication and future research directions.

2. DATA

This study used the roadside axle surveys conducted on intercity highways. These surveys have been regularly conducted to determine characteristics of heavy vehicle and freight movements in Turkey evaluated by TGDH. These surveys included weighing of trucks, randomly stopped according to a predetermined sampling ratio, in addition to interviews collecting information about the truck type (rigid or articulated), age of truck, commodity type, empty weight, load carrying capacity and total weight, as well as origin and destination provinces and country of the trip. Using the information of the load carrying capacity, empty weight and total weight, it was possible to determine the load factor (capacity utilization) of trucks. Here, it should be noted that as these surveys were performed on inter-city roads, they were capable of capturing mostly inter-city truck movements, and did not provide insight on the intra-city movements; thus, they are excluded from the scope of this study.

While a more elaborate discussion of the road side axle surveys was provided by Ozen [13], reports evaluating truck freight transportation by TGDH covering the periods of 2007-2009 and 2010-2014 have been published [14,15]. However, aforementioned specific data requirements of the model used in this study enabled the use of data only from 2007 to 2011, which is one of the reasons between small differences in selected measures also determined in this study. While this caused slight differences in measures, another reason is the data cleanup process followed in this study to eliminate surveys with missing data.

Table 1 - Descriptive statistics for model variables

Variable	Variable description	Average
LF	Load factor for loaded trips (Load factor for all trips)	78% (58%)
L	1 if a truck is loaded and 0 otherwise	69.1%
Distance	Average distance of the trip (in km)	435
Size	Average gross vehicle weight (tons) (Empty weight + load carrying capacity)	29.0
MC	Maximum allowable limit (tons)	17.9
CW	Average commodity weight (tons)	10.4
Age	Average age of vehicle (years)	6.7
Rigid	Rigid trucks	68%
Articulated	Articulated trucks	32%
Net-importer*	1 for trips made towards a net importing region; 0 otherwise	11%
Voluminous**	1 for voluminous cargo; 0 otherwise	39%

Source: Turkish General Directorate of Highways (TGDH): Road Side Axle Surveys from 2007 to 2011.

*Net Importer: regions in Turkey whose imports (in terms of the goods) is higher than exports.

**Voluminous: Goods which need an extra amount of space (e.g., timber, grains etc.).

Descriptive statistics about the data from 53,327 intercity trucks used in this study are summarized in Table 1. While 69.1% of the surveyed trucks were loaded, load factor on average was determined as 78% with an average commodity weight of 10.4 tons. While average trip distance was calculated as 435 km for the study data, it was determined by assigning the trip on the time-based shortest path between the stated origin and destination, as established by Ozen [13]. It should be noted here that this may be a partial reason causing trip length deviating slightly from the statistics published by the TGDH.

Leaving empty trucks aside constituting approximately one-third of the sample, further analysis by load type of the laden trips showed that while some load types were overrepresented others were not represented enough either due to their smaller market shares or sampling process during the axle load surveys which were performed in short sampling periods (such as 2-3 days at a study location). Thus, the current 20-commodity types by TGDH following to “Standard Goods Classification for Transport Statistics (NST)-2007” was further simplified into 6 categories as shown in Table 2, similar to the NST categories employed by TGDH prior to 2007 [16]. In this categorization Commodity Type 5 (including coke, refined petroleum, etc.) ranked first with a share of 18.3% of the sample, followed by Commodity Type 3 (food products, beverages and tobacco) with a share of 12.7 % (see Table 2). Commodity Type 1 (products of agriculture, hunting and forestry) with 10.8% share was followed by 7.9% share of Commodity Type 2. This simpler categorization enabled generation of dummy variables for load types to be used in the models. Note: Grouped goods or unidentifiable goods by NST-2007 (NST Group 18 to 20) were omitted as well as commodities that had really small shares which could not be aggregated under the selected categories in Table 2, such as, secondary raw materials, municipal wastes and other wastes (NST Group 14); mail, parcels (NST Group 15) and goods moved in the course of household and office removals (NST Group 17).

Table 2 - Classification of commodity groups

Commodity	Commodity Type	Frequency	Percent
0	Empty	16388	30.8
1	Products of agriculture, hunting, and forestry	5735	10.8
2	Coal and lignite; peat; crude petroleum Metal ores and other mining products	5328	10.0
3	Food products, beverages and tobacco	6763	12.7
4	Textiles and textile products Wood and products of wood and cork Furniture, other manufactured goods	4268	8.0
5	Coke, refined petroleum products Chemicals, chemical products Other non-metallic mineral products	9753	18.3
6	Basic metals; fabricated metal products Machinery and equipment n.e.c. Transport equipment Equip. and materials utilized in the transport of goods	4992	9.4

Source: Turkish General Directorate of Highways (TGDH): Road Side Axle Surveys from 2007 to 2011.
Number of observations: all trips (53,327); loaded trips (36,838) and empty trips (16,489).

Figure 1 and 2 show the market share of loaded (not empty, L=1) and empty (L=0) trips as a function of distance and size. Approximately 71% of the articulated trucks were loaded when surveyed, and almost half of those were traveling more than the average hauling distance of 435 km. The empty running by the articulated trucks followed the same trend, although on a smaller scale. When the rigid trucks were analyzed, a significant portion of them (loaded or not) were observed on rather short hauls of less than 400 km. More importantly, the number of trucks running empty seemed to fade out quite quickly. It is also worth noting that, for trips longer than 1000 km, distance seems to matters more for rigid trucks than for articulated trucks.

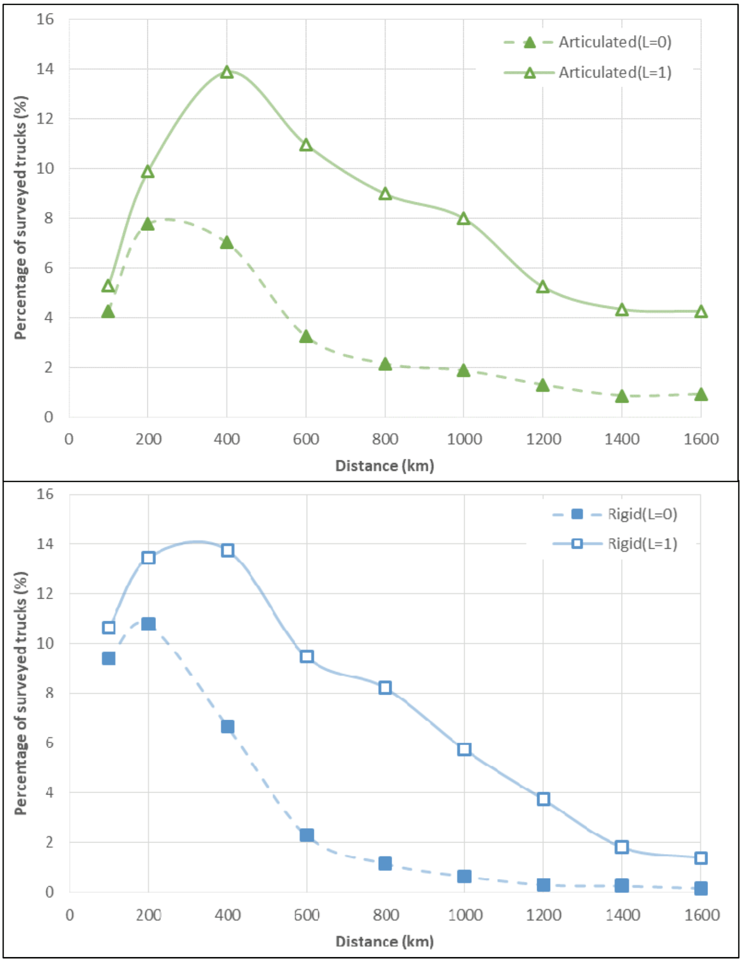


Figure 1 - Market share of loaded and empty trips as a function of distance

The distribution of road freight movements surveyed during 2007 to 2011 by truck size (see Figure 2) supports the fact that articulated and rigid trucks have different characteristics. In this study, truck size is defined in Gross Vehicle Weight (GVW), which is the sum of maximum load carrying capacity and empty weight of the vehicle. As the size of the articulated trucks increases, the percentage of loaded trips within the size categories increases, hence the percentage of empty runs decrease. For rigid trucks, the percentage of loaded trips within each truck size category fluctuates without suggesting a clear trend as is the case for articulated trucks.

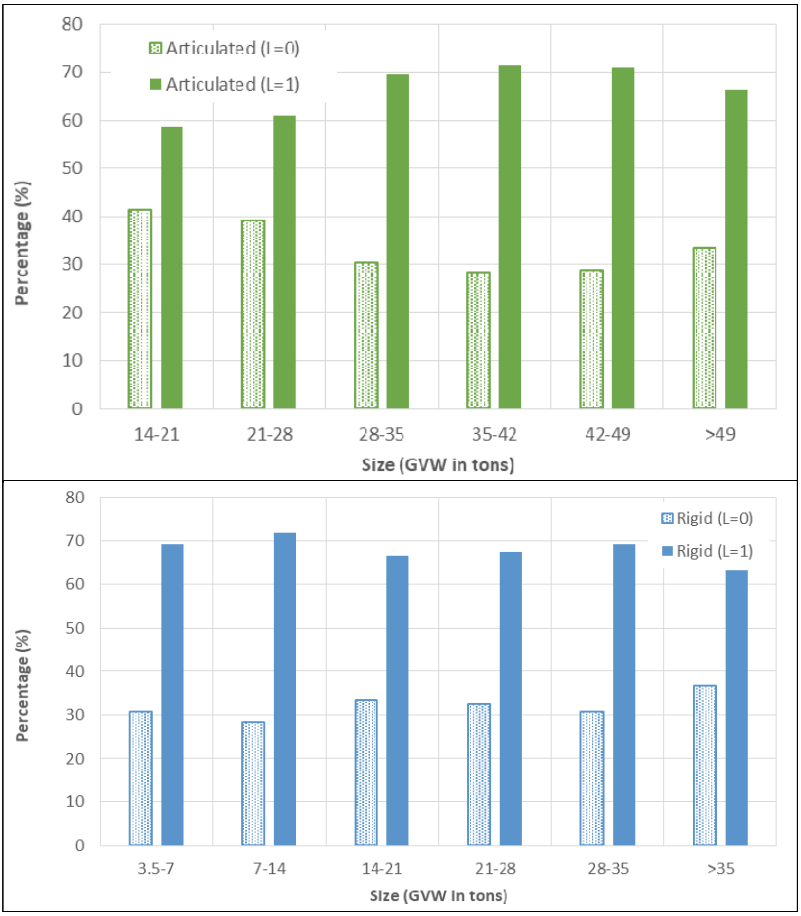


Figure 2 - Market share of loaded and empty trips as a function of truck size (Gross Vehicle Weight)

2.1. Limitations

One of the limitations in the data is that the surveys are conducted on state roads, which include mainly intercity freight transportation. As a result, the analysis is not representative for intra-city freight transport, which will behave differently. Secondly, information about trip chains, truck tours, and warehouses, as well as loading and unloading at transitional hubs, is not included in our data set. This would have made possible performing a longitudinal study. Lastly, unavailability of driver's salary related data. One major difference from developed country as the cost of the truck, although with a comparatively older age, and fuel cost is not dramatically different but the driver salary is. This could explain issues like the readiness to wait for a balancing load and the labor that can be rationally spent on loading the truck carefully. However, it should be noted that the data set is more disaggregated than is usual for a developing country. The data is almost similar to the Danish data set used in Abate [10] and to data in other EU countries. But, as mentioned above, more recent data (data after 2011) did not provide required details preventing their use in the modeling step of this study. Nevertheless, for many of the developing countries, data are hard to obtain due to privacy reasons; most importantly, in the case of developing countries, such data are almost never available due to lack of surveys and resources. However, beyond roadside survey data, it is necessary to collect commodity flow data at a national scale as in the USA.

3. METHODOLOGY

3.1. Factors Affecting the Load Factor

The load factor can be defined as the ratio of the commodity weight to the maximum carrying capacity of the truck, expressed in percentage. To evaluate capacity utilization with respect to the load factor, weight and volume (density) of the respective commodities should ideally be taken into account. However, such density data are often not available (as in this study). As a result, this study applied a weight-based load factor percentage (LF_i). This is defined as the ratio between commodity weight (CW_i) and maximum carrying capacity (MC_i).

$$LF_i(\%) = 100x\frac{CW_i}{MC_i} \quad (1)$$

Factors affecting LF_i include distance, commodity groups (to account for heterogeneity of potential high density cargo versus low density cargo), and truck size.

3.2. Factors Affecting Market Access

The market access decision refers to the probability that the truck is loaded rather than empty. Generally, as the distance increases, the percentage of empty runs is expected to go down. It is not clear why there is a linear relationship between the distance and the probability that the truck is loaded. It may well be that there is a marginal decreasing effect of distance. To account for this, a Box-Cox model is also worked out. In fact, it is revealed that the Turkish transport market essentially constitutes two quite different transport markets, one for articulated trucks and another for rigid trucks. For articulated trucks, which are generally larger, the probability that the truck is loaded increases with size. The reverse is true for rigid trucks because rigid trucks are relatively small. Market access decisions may also depend on the freight movement balance, i.e., whether the region is a net-importing or net-exporting region [10]. The probability that the truck is loaded is usually higher for trips towards a net-importing region because it is not easy to find matching return trips. Conversely, because of a higher probability of matching return load from net-exporting regions empty trucks are more likely to be observed towards such regions [9]. Net exporting and net importing zones have been classified from overall freight movement in Turkey and this paper has introduced a dummy variable indicating whether the trip was made to a net-importing or net-exporting region. Furthermore, truck traffic to and from the very large provinces, i.e., İstanbul, Ankara, and İzmir, will tend to have different characteristics, which in turn may affect the probability of loading. To capture this effect, another dummy variable was introduced for traffic going to and from these large regions. The age of the truck is usually assumed to be negatively related to the market access decision. It should be noted that, for Turkey, the truck age profile is quite different from those in the EU and other developed countries. The average age of trucks is higher and there is a relatively large segment of trucks older than 12-14 years, most of which are rigid. Hence, in some cases, the market access decision might not be altered by the age of the truck, particularly for rigid trucks.

3.3. Model Specification

Truck utilization can be modeled as a sample selection problem because the dependent variable (the load factor) is observed only when the truck is loaded. This was proposed by Abate [10], applying the standard Heckman model [12] to the decision to drive with the truck loaded or empty (henceforth referred to as the market access decision) and the load-factor percentage conditional on the truck carrying a load. Clearly, companies and freight carriers prefer to minimize empty running as much as possible, while also trying to maximize the load factor for loaded trucks to maximize profits.

The decision to load the truck is essentially connected to the market access of the carrier and may relate to characteristics of the carrier, the trip and the type of truck. The load-factor percentage will largely depend on the same variables; however, for identification purposes, it is required that some variables are used as exclusion restrictions in the market access decision equation but not in the load factor. If we let LF_i represent the load-factor percentage for observation i and L_i the corresponding binary choice of loading the truck, the model can be written as:

$$L_i^* = f(X_{1i}, \theta) + \upsilon_i \tag{2}$$

$$L_i = \begin{cases} 1, & if\ L_i^* > 0 \\ 0, & if\ L_i^* \leq 0 \end{cases} \tag{3}$$

$$LF_i^* = g(X_{2i}, \beta) + \varepsilon_i\,,\ \text{if}\ L_i = 1 \tag{4}$$

$$LF_i = \begin{cases} LF_i^*, & if\ LF_i^* > 0 \\ 0, & if\ LF_i^* \leq 0 \end{cases} \tag{5}$$

X_{1i} and X_{2i}, respectively, represent vectors of explanatory variables for the binary choice of L_i and the load-factor percentage equation (LF_i); θ and β represent the structural parameters. A loaded truck is represented by L_i=1. X_{2i} contains all the variables in X_{1i} and additional variables for identification (exclusion restrictions). The residuals υ_i and ε_i are jointly normal with zero mean. The assumptions for estimation are as follows [17]:

1) (X_{2i}, L_i) are always observed, but LF_i is only observed when L_i=1 (sample selection);

2) (υ_i, ϵ_i) is independent of X_{2i} with zero mean;

3) $\upsilon_i \sim$ normal (0,1) (distributional assumption); and

4) $E(\upsilon_i|\epsilon_i) = \gamma_2\epsilon_i$ (residuals are jointly normal and may be correlated).

In addition to these assumptions, it is necessary to remember that the LF_i distribution is essentially truncated at 0 (cannot have a negative load factor). Hence, a truncated version of the Heckman model is considered (see Eq. 3-4). It should be said, however, that considering the truncated version of the LF_i distribution makes relatively little difference in this example, and the problem of considering an un-truncated version seems to be relatively small [10].

The simplest way of estimating the model would be to assume that the load-factor percentage could be modeled by considering only the loaded trips. However, as documented by Heckman [12] and others, this will lead to a sample selection problem, which in turn leads to biased estimates. For instance, it is possible that trucks opt to run empty instead of loaded because the cost of running loaded exceeds the revenue [9]. A joint estimation is needed to model the load factor to understand its determinants at a population level. Therefore, a joint estimation model developed by Heckman implemented in this study, which has been widely used in many areas of science including transportation [18-20]. A more elaborate discussion of Heckman model was provided by de Jong [19]. As a result, this paper estimates the parameters of Eq. 1-4 by full information maximum likelihood (FIML), to account for joint correlation between the two endogenous components, using the QLIM procedure in the SAS Software. The complete likelihood function is given by:

$$\begin{aligned} LL(\beta,\theta\lambda) = &\sum_{i|L=0} ln\left(1-\Phi\big(f(X_{1i},\theta)\big)\right) \\ &+\sum_{i|L=1}\Bigg[ln\left(\phi\left(\frac{LF_i-g(X_{2i},\beta)}{\sigma}\right)\right)-ln\sigma \\ &+ln\left[\Phi\left(\frac{f(X_{1i},\theta)+\rho\left(\frac{LF_i-g(X_{2i},\beta)}{\sigma}\right)}{\sqrt{1-\rho^2}}\right)\right] \\ &+ln\left(\Phi\left(\frac{f(X_{1i},\theta)}{\sigma}\right)\right)\Bigg] \end{aligned} \tag{6}$$

According to Eq. 6, the first part of the function represents the censored part, where $L_i = 0$. The un-censored part consists of the equation resulting from the LF_i model (adjusted for truncation) and the choice of L_i. The model shown in Eq. 6 with linear-in-parameter specifications for f and g is considered a base model and is referred to as Model 0. The base model do not differentiate between types of trucks but estimate the Heckman model [12] for the whole sample, which reflects the model structure applied in Abate [10]. However, different truck types could behave differently in regard to the load-factor percentage and the share of empty runs, therefore, two alternative versions of Model 1 are estimated, one for rigid trucks and one for articulated trucks, in order to measure differences between the types of trucks. Finally, these two models are formulated in more comprehensive versions, in Model 2, where main modifications are introduced.

Firstly, Model 2 allows for a richer parameterization of the distance and size variables in the LF model in order to account for differences in how commodity groups are described. This turns out to be relatively important as there are big differences in how size and distance affect the load-factors percentage for different commodities. Secondly, for the LF_i model, Model 2 account for unobserved heteroscedasticity in the error-term ε_i by representing the variance as a function of the distance. Hence, the variance in the LF_i model is now given by;

$$E(\varepsilon_i^2) = \sigma_i^2 = \sigma_i^2\left(1 + \gamma ln(Dist_i)\right) \tag{7}$$

Generally, it is important to account for heteroscedasticity because the standard Heckman model can produce unreliable results if not accounted for, as shown in Arabmazar and Schmidt [21]. The robustness of the estimators for the sample selection models has been taken into account by [22-23]. Finally, Model 2 relax the linear-in-parameter assumption for the L_i model by estimating a Box-Cox model for the distance variable. Hence,

$$BC(Dist_i) = \begin{cases} \frac{Dist_i^{\lambda}-1}{\lambda} & if\ \lambda \neq 0 \\ ln(dist_i) & if\ \lambda = 0 \end{cases} \tag{8}$$

The linear-in-parameter specification, which is a special case of the Box-Cox parameter [24-26] when $\lambda = 1$, is clearly rejected.

4. RESULTS

Table 3 and 4 presents the main results for the two different model segments for the rigid and articulated truck. Note that *t*-values at vehicle levels are included in parenthesis. As mentioned above, Model 0 in Table 3 and 4 is the simple base model without segmentation, whereas Model 1 includes one model for rigid and one for articulated trucks. Hence, the assumption of identical parameters for different truck types across the sample is relaxed. Model 2 includes the three additional econometric improvements. For the market access equation, three exclusion restrictions represented by age, net-import (freight movement balance), and dummies for large provinces are included (see Table 4). These variables have significant effects on the market access model, with empty running negatively related to age and positively related to the “net-import” variable and the large provinces dummies. On the other hand, in Model 2, the effect of age is positively significant but almost equal to zero. There are two reasons. First, there are more rigid trucks than articulated trucks in Turkey. Secondly, in this data-set, trucks older than 15 years are 10 percent of the whole rigid truck fleet, while trucks older than 10 years correspond to 28 percent, which is a substantial proportion and can’t be ignored. Although the new road freight transport regulations in 2003 favor moving freight by articulated trucks, it will take a long time to replace older rigid trucks with newer articulated trucks. Besides, the overall age distribution of heavy goods vehicles in Turkey is older compared to the EU or other developed countries; this is a common trend in developing countries. The significance of the correlation coefficient between the residual terms, ρ, implies that the joint estimation is appropriate and that the sub-sample of loaded trips is not selected randomly.

Table 3 - FIML Heckman Model estimates for LF equation

	Combined	Rigid	Articulated	Rigid	Articulated
Variables	Model 0	Model 1	Model 1	Model 2	Model 2
Distance	-0.0097 (-11.43)	-0.0098 (-7.18)	-0.0117 (-11.54)		
Distance 1 (Distance×Commodity group 1)				0.0089 (4.35)	0.0043 (1.22)
Distance 2 (Distance×Commodity group 2)				0.0059 (1.53)	0.0161 (4.78)
Distance 3 (Distance×Commodity group 3)				0.0205 (12.62)	0.0068 (2.49)
Distance 4 (Distance×Commodity group 4)				0.0305 (11.79)	0.0152 (4.81)
Distance 5 (Distance×Commodity group 5)				0.0123 (5.66)	-0.0018 (-1.02)
Distance 6 (Distance×Commodity group 6)				0.0216 (8.71)	0.0044 (2.09)
Size	-0.0020 (-6.84)	-0.0030 (-6.51)	-0.0339 (-37.84)		
Size 1 (Size×Commodity group 1)				0.0036 (6.11)	-0.0001 (0.13)
Size 2 (Size×Commodity group 2)				0.0088 (16.08)	0.0003 (0.68)
Size 3 (Size×Commodity group 3)				0.0019 (3.49)	-0.0006 (-1.01)
Size 4 (Size×Commodity group 4)				-0.0057 (-7.44)	-0.0076 (-8.99)
Size 5 (Size×Commodity group 5)				0.0041 (9.09)	0.0010 (2.34)
Size 6 (Size×Commodity group 6)				-0.0054 (-7.95)	-0.0040 (-6.94)

Table 3 - FIML Heckman Model estimates for LF equation (continue)

	Combined	Rigid	Articulated	Rigid	Articulated
Variables	Model 0	Model 1	Model 1	Model 2	Model 2
Voluminous	-0.1290 (-22.58)	-0.1378 (-18.57)	-0.1135 (-14.35)	-0.0514 (-7.95)	-0.0742 (-7.77)
Constant	1.0224 (86.69)	1.0250 (59.25)	2.3124 (61.65)	0.5285 (63.33)	0.7482 (67.07)
Rho	-0.0330 (-1.20)	-0.0129 (-0.31)	-0.0419 (-0.76)	0.9706 (530.29)	0.5556 (87.67)
Sigma	0.4764 (240.17)	0.5071 (196.31)	0.3748 (135.84)	1.0509 (91.88)	0.8227 (46.02)
Hetero Log. Distance				-0.1149 (-128.62)	-0.1097 (-62.54)
No. of observations	53327	36321	17006	36321	17006

Table 4 - FIML Heckman Model estimates for L (market access decision) equation

	Combined	Rigid	Articulated	Rigid	Articulated
Variables	Model 0	Model 1	Model 1	Model 2	Model 2
Distance	0.0861 (38.70)	0.1101 (34.76)	0.0590 (18.54)	0.1102 (12.47)	0.1420 (7.13)
Size	-0.0082 (-10.38)	-0.0084 (-7.60)	0.0008 (0.29)	-0.0038 (-6.39)	0.0191 (8.12)
Age	-0.0134 (-9.45)	-0.0101 (-5.96)	-0.0242 (-8.62)	0.0020 (2.48)	-0.0038 (-1.80)
Net-Import	0.2286 (9.37)	0.2115 (7.15)	0.2737 (6.34)	0.0849 (4.79)	0.1909 (5.20)
Large Provinces	0.3204 (18.32)	0.3641 (16.23)	0.2014 (6.92)	0.1355 (9.41)	0.2370 (10.17)
Lambda2.L (distance)				0.7262 (12.68)	0.4590 (4.97)
Constant	0.5643 (20.07)	0.4688 (13.75)	0.4138 (3.62)	0.4689 (26.36)	-0.4114 (-4.37)
No. of observations	53327	36321	17006	36321	17006

In Models 0 and 1 in Table 3, trip distance appears to be negatively associated with the load factor, i.e., if the trip distance increases, the load factor decreases for both rigid and articulated trucks. This finding does not comply with earlier findings in the literature. However, for the improved Model 2, which accounts for unobserved heterogeneity and expand the parameterization of distance and size, the signs for all distance variables become positive (except for commodity group 5 for the articulated trucks but is not significant). There are significant differences between commodity groups. For commodity groups 3, 4 and 6, distance tends to be quite significant. This fits well with the fact that these commodities are usually among the most "flexible" in the sense that they do not require special equipment or handling. Hence, there is a higher chance of an increase in the overall load factor. Furthermore, the voluminous dummy variable, which indicates whether or not a commodity is voluminous, tends to reduce the load-factor. It is clear from that distance matters more for rigid trucks than for articulated trucks. Furthermore, the negative and highly significant values of Hetero Log distance imply that there is higher variability in the load factor for short distances, whereas the variability in the load decreases for longer distances.

The effect of trip distance is positive and significant on the market access decision in all models, indicating that loaded trucks are preferred for longer distance, irrespective of the type of truck. Size has a negative effect on the load factor in model 1 for both rigid and articulated trucks, which is similar to the finding of Abate [10]. However, when truck size was further classified according to the commodity groups in Model 2, it appears that size has a significantly negative effect on the load factor for commodity groups 4 and 6. This implies that increase in size leads to a decrease in the load factor. The interpretation is that these commodities are difficult to merge and load into a larger truck, while it is easy to fill the smaller trucks with such goods. On the contrary, commodity group 2 is very easy to collect and compact in the form of a concrete mass, and hence for larger trucks, it is much easier to combine this cargo and hence increase the overall load factor. The effect of size on the market access decision is different for rigid and articulated trucks. It has been shown that size has a negative and significant effect for rigid trucks (in model 1 and 2), but a positive and significant effect for articulated trucks (in model 2). This replicates the results of Abate [10], where size has a negative effect on the market access decision, while size-squared has a positive effect; this implies that carriers prefer high capacity trucks to be loaded.

Mostly, studies based on selection models find positive signs for "rho". Although in Model 0 and 1 rho sign is negative, in Model 2 where three extensions were introduced the rho sign has become positive and quite significant; infers the credibility of Model 2. A detailed explanation is given in Abate [10] for why one might get a positive or negative sign. The positive rho sign might occur if unobserved effects that make a vehicle more likely to be loaded also cause the vehicle to have a higher load factor. A negative sign implies the opposite.

Elasticity for selected variables was estimated (see Table 5 and Figure 3), for the market access equation for Model 1 and Model 2. The elasticities show the proportionate change in the probability that a truck is loaded, for a proportionate change in the continuous explanatory variables. The effect of the distance is positive, i.e., the probability that the truck is loaded increases if the distance increases. For example, a 1 percent increase in distance will increase the probability that the truck is loaded by 0.13% for rigid trucks in Model 1. The elasticity estimates of size show that for rigid trucks, the probability that the truck is loaded increases if the size decreases; the opposite is true in the case of articulated trucks. However, in Model 1, the effect of size is not significant for the articulated trucks. In Model 2, for example, a 1% increase in the size of rigid trucks decreases the probability that the truck is loaded by 0.17%, while a 1% increase in the size of articulated trucks increases the probability that the truck is loaded by 0.68%. For articulated trucks, the elasticity effect of the age of the truck is such that an increase in age will lead to more empty runs. This is in line with expectations. For the articulated trucks in Model 2, however, the effect of age is not significant.

Table 5 - Elasticity estimates based on truck type for Market Access Equation (L)

	Model 1		Model 2	
Variables	Rigid	Articulated	Rigid	Articulated
Distance	0.1687	0.1286	0.1518	0.1381
Size	-0.1452	0.0087	-0.1713	0.6552
Age	-0.0852	-0.0894	0.0357	-0.0764

Note: The elasticities are calculated by the Delta method. The elasticities show percentage changes in the probability that the truck is loaded or empty for a proportionate change in the explanatory variables.

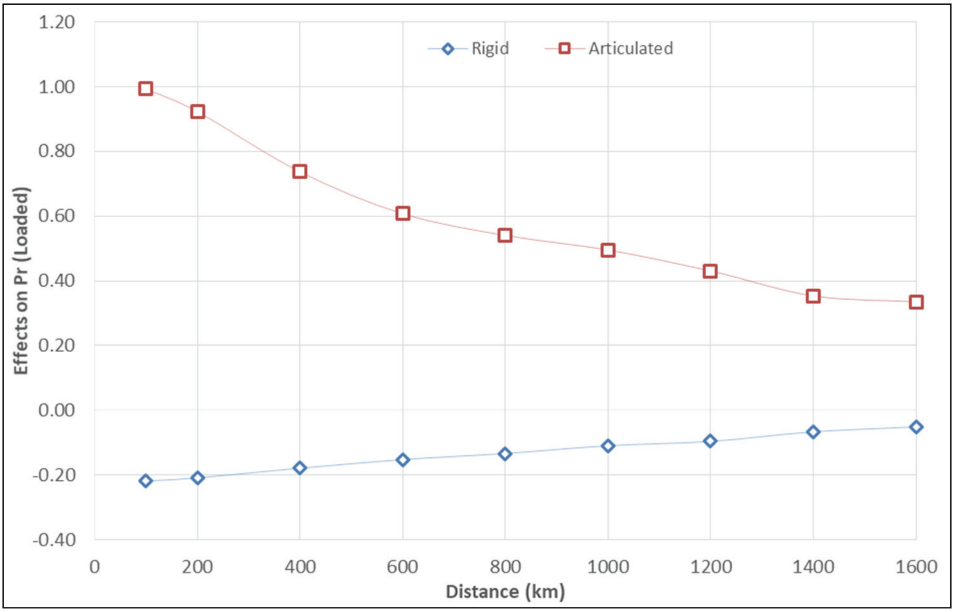


Figure 3 - Average marginal effects of truck size on the probability that the truck is loaded based on Model 2, shown as a function of distance

4.1. Validation

Figure 4 and 5 presents a validation of Model 2. To do this, the modeled probability distribution that the truck is loaded against distance and size is plotted, and compared this with the observed distributions (in the survey). It is seen that model closely predicted the observed distributions. As discussed above, the average number of loaded trips increases with increasing trip length. This is the case for rigid as well as for articulated trucks. The corresponding probability distribution with respect to the size of the truck shows an interesting pattern. For articulated trucks, the probability that the truck is loaded increases with size, whereas it is the other way round for rigid trucks. Figure 4 and 5 underlines the fact that the transport markets for rigid and articulated trucks are widely different.

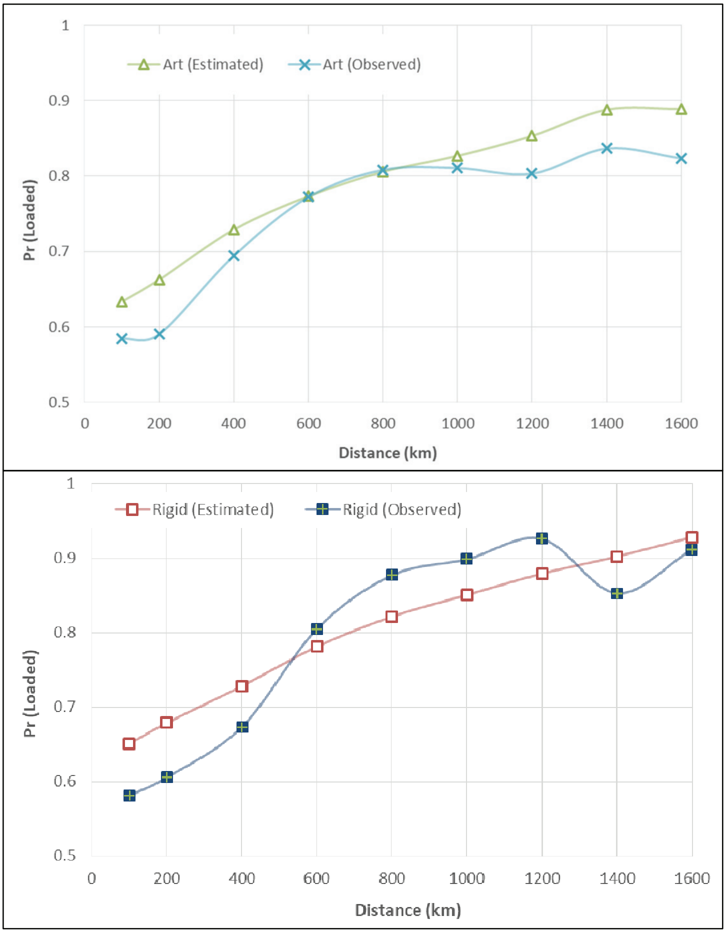


*Figure **4** - Probability that the truck is loaded as a function of distance*

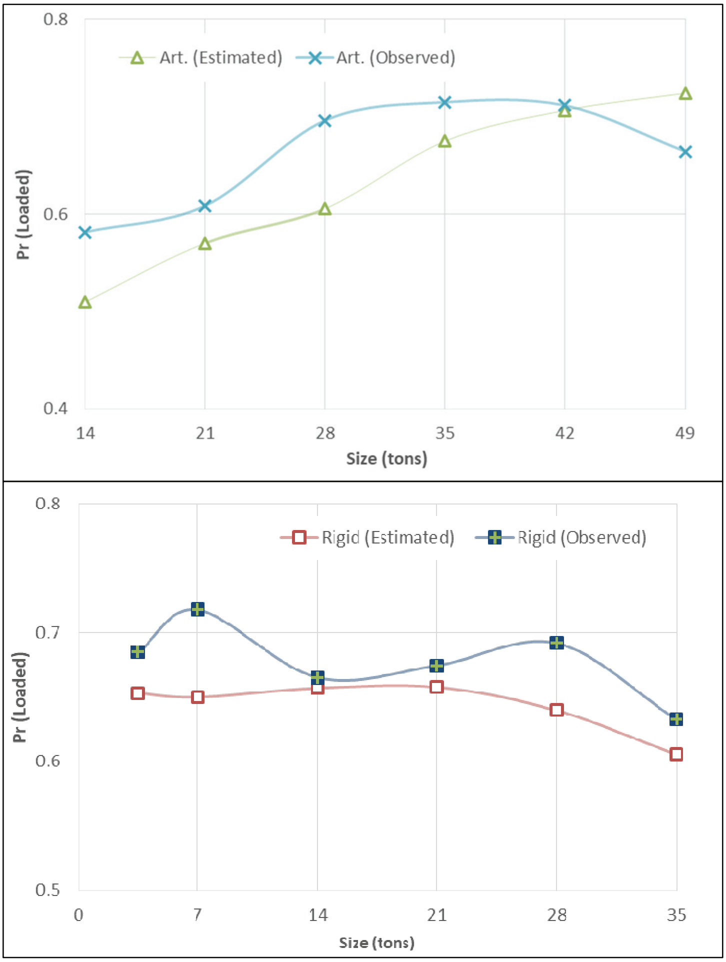


Figure 5 - Probability that the truck is loaded as a function of truck size

5. CONCLUSION AND DISCUSSION

Road freight transport corresponds to 89% of overall freight transportation in Turkey, which is very imbalanced, but is not expected to change significantly for many years to come, as alternative modes are not developed properly. This obviously increases the importance of efficiency considerations for the trucking industry because this is a main entry point to change the environmental footprint and economic impact in the country as a whole. The objective of this study was to investigate the main factors affecting capacity utilization in truck freight transport, in Turkey. The current paper extends the work by Abate [10] in three ways. Firstly, this study has introduced econometric extensions by partly relaxing the assumption of linearity of distance effects in the empty run model and by relaxing the assumption of homogeneity in the specification of the variance in the load factor model. Secondly, this paper has presented a more explicit representation of commodity groups for which we estimate separate parameters with respect to distance and size effects in the load-factor model. Thirdly, this study considers different models for rigid and articulated trucks, which, in turn, are found to be important.

The results suggest that trip distance positively influences both the load factor and the probability of loaded trips. In the baseline model, the size of the truck was shown to have a negative relationship with the load factor and the probability that the truck is loaded. However, in a more advanced model where models are segmented according to the type of truck, it was shown that rigid and articulated trucks behave differently. Hence, the probability that an articulated truck is loaded is positively affected by size, whereas the opposite is true for rigid trucks. The fact that these two types represent essentially different transport markets is interesting for Turkey as well as other developing countries where the same mix of vehicles exist. It may well signal that transport policies aimed at improving efficiency for developing countries should be able to differentiate between the type of truck and possibly even be targeted toward different weight categories. Another important finding is the importance of the commodity type. As it can be seen in the modeling results, the capacity utilization pattern is quite different between commodity groups when combined with size and distance separately. While some commodity types resulted in a net positive relation for size, others had negative ones suggesting that probability of being loaded was higher for them. Thus, any further plans focusing on minimization of the empty or inefficient movements on highways must be developed considering commodity types and their flow characteristics on the intercity corridors in Turkey.

Further Recommendations

Although the paper is a step toward better understanding capacity utilization in the trucking industry, the current findings may have some generic value. For instance, the findings about how truck size influences capacity utilization may have implications for legislation related to allowing mega trucks. Hence, as the size of trucks increases, it may be possible to derive indicative measures on how this will impact the utilization of capacity. Secondly, increase in loading probability by transport distance is a simple reflection of increase in total cost, if truck is empty, which is simply pressurizing drivers to find load for longer trips, while they can risk traveling empty on shorter distances. A complete understanding of capacity utilization for road transport requires further research. Most importantly, more data is needed, as described below:

- A correct representation of freight density and value density is important.
- A better description of the freight operator in order to take account of the heterogeneity between small and large companies is necessary.
- A more detailed representation of logistic activities is needed. Such data includes information about delivery routing, trip chains, warehouses, loading and unloading at transitional locations, as well as restrictions in terms of time windows and dependence on other operators and modes. These hidden constraints may potentially influence and reduce the potential for capacity utilization.

- A more detailed monitoring of Turkish trucks is needed, including intra-regional trips, and not just trips between regions.

Thus, it is necessary to collect national commodity flow data (CFD) as in some developed countries. A detailed discussion of such a national freight data framework was proposed in TRB [27], with many aspects discussed more recently in detail in a workshop [28] to improve the quality of commodities to be transported and to improve commodity flow data [29]. Similarly, the need for building the origin-destination freight (and passenger) matrices for pan-European region was the focus of the project titled "Methodological Framework for Modeling European Passenger and Freight Transport (MYSTIC)", which also discussed involvement of various parties in freight transportation, such as carriers, intermodal platforms, etc. During this European Union Harmonization Process, as for a candidate country, it is also necessary to push for more integrated and complete freight data in Turkey, bringing public and private parties together to share their data.

Furthermore, freight transportation is an important factor for the development of the national economies. It is always desirable to minimize total cost of freight transportation (private and external costs), while goods are efficiently transported. Private costs are the direct expenses incurred by providers of freight transportation. Freight transportation has also costs that are paid by others. Such costs are referred as external costs (i.e., accidents, emissions, and other costs associated with the operation and maintenance of roads and bridges). If policies are adopted to charge the full cost of freight transportation, transportation users start to consider these costs when making shipping decisions. This may cause major structural changes to reduce external costs. One impact would be reduction of hauling distances by bringing production centers closer to consumption ones [30,31]. Another one would be combining inefficient movements to achieve higher level of capacity utilization in freight. However, to manage the freight sector in a more efficient and successful way, it is also necessary to invest on Intelligent Transportation Systems (ITS) in data collection as well as freight planning and management [32].

References

[1] T.C. Ulaştırma Denizcilik ve Haberleşme Bakanlığı, İstatistiklerle Ulaştırma Denizcilik ve Haberleşme: 2003-2016, Strateji Geliştirme Başkanlığı, 2016.

[2] Organisation for Economic Co-operation and Development/Intertnational Transport Forum, ITF Transport Outlook, OECD Publishing, 2017.

[3] Eurostat, Freight transport statistics, 2017. http://ec.europa.eu/eurostat/statistics-explained/index.php/Freight_transport_statistics Last Access: April 1, 2018.

[4] European Comission, White Paper on Transport: Roadmap to a Single European Transport Area: Towards a Competitive and Resource-efficient Transport System. Publications Office of the European Union, 2011.

[5] Abate, M., Kveiborg, O., Capacity Utilisation of Vehicles for Road Freight Transport, in: Ben-Akiva, M.E., Meersman, H., Voorde, E. van de (Eds.), Freight Transport Modelling. Emerald Group Publishing, Bingley, UK, 287–304, 2013.

[6] Ozen, M., Tuydes-Yaman, H. Evaluation of emission cost of inefficiency in road freight transportation in Turkey. Energy Policy, 62, 625–636, 2013.

[7] Piecyk, M. I., Analysis of Long-Term Freight Transport, Logistics and Related CO_2 Trends on a Business-as-Usual Basis. Ph.D. Thesis, Heriot-Watt University, 2010.

[8] McKinnon, A. C., Benchmarking Road Freight Transport. Review of a Government-Sponsored Programme. Benchmarking: An International Journal, 16 (5), 640 – 656, 2009.

[9] Beilock, R., Kilmer, R. L.. The Determinants of Full-Empty Truck Movements. American Journal of Agricultural Economics, 68(1), 67–76, 1986.

[10] Abate, M., Determinants of Capacity Utilisation in Road Freight Transportation. Journal of Transport Economics and Policy, 48(1), 137–152, 2014.

[11] Abate, M., de Jong, G., The Optimal Shipment Size and Truck Size Choice – The Allocation of Trucks Across Hauls. Transportation Research Part A: Policy and Practice, 59, 262–277, 2014.

[12] Heckman, J. J., Sample Selection Bias as a Specification Error. Econometrica, 47(1), 153–161, 1979.

[13] Ozen, M., Estimation of Road Freight Transportation Emissions in Turkey. Ph.D Thesis. Middle East Technical University, 2013.

[14] Turkish General Directorate of Highways, Karayollarında Ağır Taşıt Trafiğinin ve Yük Taşımacılığının Özellikleri ve Eğilimleri: 2007-2008-2009 Yılı Etüt Sonuçları. Turkish General Directorate of Highways, 2011.

[15] Turkish General Directorate of Highways, Karayollarında Ağır Taşıt Trafiğinin ve Yük Taşımacılığının Özellikleri ve Eğilimleri: 2010-2011-2012-2013-2014 Yılı Etüt Sonuçları. Turkish General Directorate of Highways, 2011.

[16] Unal, L., Modeling of Freight Transportation on Turkish Highways. Ph.D Thesis. Middle East Technical University, 2009.

[17] Wooldridge, J. M., Econometric Analysis of Cross Section and Panel Data. Cambridge, Mass: The MIT Press, 2010.

[18] Portugal-Perez, A., Wilson, J. S., Export Performance and Trade Facilitation Reform: Hard and Soft Infrastructure. World Development, 40(7), 1295-1307, 2012.

[19] de Jong, G., Tanner, R., Rich, J., Thorhauge, M., Nielsen, O. A., Bates, J., Modelling Production-Consumption Flows of Goods in Europe: The Trade Model Within Transtools3. Journal of Shipping and Trade, 2(1), 5, 2017.

[20] Tsotsotso, K., Montshiwa, E., Tirivanhu, P., Fish, T., Sibiya, S., Mlangeni, T., Moloi, M., Mahlangu, N., Determinants of Skills Demand in A State-Intervening Labour Market: The Case Oof South African Transport Sector. Higher Education, Skills and Work-Based Learning, 7(4), 408-422, 2017.

[21] Arabmazar, A., Schmidt, P., An Investigation of the Robustness of the Tobit Estimator to Non-Normality. Econometrica, 50(4), 1055–1063, 1982.

[22] Maddala, G. S., Nelson, F. D., Specification Errors in Limited Dependent Variable Models. NBER Working Paper No: 0096. National Bureau of Economic Research, Inc., 1975.

[23] Nelson, F. D., The Effect of and a Test for Misspecification in the Censored-Normal Model. Working Paper No: 291. California Institute of Technology, Division of the Humanities and Social Sciences, 1979.

[24] Hurd, M., Estimation in Truncated Samples when There is Heteroscedasticity. Journal of Econometrics, 11(2–3), 247–258, 1979.

[25] Sakia, R. M., The Box-Cox Transformation Technique: A Review. Journal of the Royal Statistical Society: Series D (The Statistician), 41(2), 169–178, 1992.

[26] Box, G. E. P., Cox, D. R., An Analysis of Transformations. Journal of the Royal Statistical Society: Series B (Methodological), 26(2), 211–252, 1964.

[27] Transportation Research Board of the National Academies, Special Report 276: A Concept for a National Freight Data Program. Transportation Research Board of the National Academies, 2003.

[28] Transportation Research Board of the National Academies, Transportation Research Circular: Commodity Flow Survey Conference. Transportation Research Board of the National Academies, 2006.

[29] US Census, Commodity Flow Survey, 2015. https://www.census.gov/programs-surveys/cfs/data/datasets.html Last Access: September 20, 2018.

[30] Forkenbrock, D. J., External Costs of Intercity Truck Freight Transportation. Transportation Research Part A: Policy and Practice, 33, 505–526, 1999.

[31] Forkenbrock, D. J., Comparison of External Costs of Rail and Truck Freight Transportation. Transportation Research Part A: Policy and Practice, 35, 321–337, 2001.

[32] Crainic, T. G., Gendreau, M., Potvin, J. Y., Intelligent Freight-Transportation Systems: Assessment and the Contribution of Operations Research. Transportation Research Part C: Emerging Technologies, 17(6), 541–557, 1999.

Teknik Dergi, 2020 9833-9850, Paper 569

Adoption of Virtual Reality (VR) for Site Layout Optimization of Construction Projects

Abbas Ahmad MUHAMMAD[1]
Ibrahim YITMEN[2]
Sepehr ALIZADEHSALEHI[3]
Tolga CELIK[4]

ABSTRACT

Among the significant tasks of construction site management, site layout planning is a crucial activity that influences project success. Currently, the concept of an optimised construction site layout emerges as a result of executing tools and techniques that the construction project management discipline provides. Despite the plentiful existence of site layout techniques, the authors of this paper believe that the incorporation of virtual reality (VR) for jobsite organisation is a valuable and modern management phenomenon for optimal site layout planning. Therefore, the purpose of this paper is to assess traditional practices for site layout planning and introduce the application of VR in construction jobsite organisations for site layout planning, collision detection, and evaluation of construction site layout scenarios. In this paper, two different construction jobsite scenarios for the structural phase of a construction are created and for optimisation purposes, Autodesk Revit software, SketchUp, and Lumion are used to develop the three-dimensional (3D) model. A VR headset is used for testing purposes of the proposed 3D jobsite organisation model. A descriptive case study is examined to describe the proposed organisational plan for the scenarios. In this study, a questionnaire survey has been chosen as one of the research instruments for data collection and evaluation purposes. It is found out that traditional 2D methods used in jobsite layout planning is easier to understand and less time consuming for users compared to 3D VR based jobsite layout organisation. On the other hand, results indicated that 3D VR based jobsite layout planning is more effective to comprehend by users and enhances the ability of collision detection. The results presented in this paper aim to provide industry professionals with a

Note:
- This paper has been received on May 14, 2018 and accepted for publication by the Editorial Board on March 20, 2019.
- Discussions on this paper will be accepted by May 31, 2020.
- https://dx.doi.org/10.18400/tekderg.423448

1 Eastern Mediterranean University, Civil Engineering Department, Gazimagusa, Cyprus - ahmadabbas430@yahoo.com - https://orcid.org/0000-0003-0758-6414

2 Jönköping University, Department of Construction Engineering and Lighting Science, Jönköping, Sweden - ibrahim.yitmen@ju.se - https://orcid.org/0000-0003-4288-9904

3 Northwestern University, Department of Civil and Environmental Engineering, Illinois, USA - sepehralizadehsalehi2018@u.northwestern.edu - https://orcid.org/0000-0003-1567-2590

4 Eastern Mediterranean University, Civil Engineering Department, Gazimagusa, Cyprus - tolga.celik@emu.edu.tr - https://orcid.org/0000-0002-2513-6735

better understanding of the benefits and advantages of how VR technology can improve the effectiveness of jobsite layout planning compared to traditional practices.

Keywords: Virtual reality, site layout plan, construction industry, jobsite organisation, 3D model.

1. INTRODUCTION

Regardless of the type and size of the construction project, jobsite layout planning is developed subsequent to a contractor being awarded the bid and prior to the contractor's receipt of the notice to proceed [1]. Planning the site layout of construction projects is a critical task that can significantly affect productivity and efficiency, and hence, the success of the project [2]. Site layout planning is a complex task in the majority of construction projects because of the diversity of decision variables, including but not limited to identification, sizing placing, and coordination of temporary facilities, plants, materials, and people within the boundaries of a construction site [3]. However, due to the heterogeneity of construction activities, that is in the nature of its organisation – project design, time constraints, work areas, environmental conditions, and facility locations, etc. – site layout plans turn out to be unique for each construction project. As the site layout plans are prepared to ensure that work within the construction site progresses smoothly without any obstruction to obtain optimal use of resources and space, preparation of these plans is usually the responsibility of contractors awarded with the job. An effective and systematic approach for planning an optimised site layout plan ensures the optimal use of available workspace, leading to cost and time savings during the construction process, promoting a safe working environment, and providing safe access into and out of the site [2, 4].

According to Ning et al. [5], construction site layout planning has been perceived as a critical step in construction planning by experts and scientists and it is considered to be a decision-making process that includes recognising issues and opportunities, creating solutions, and picking the best option and actualising it. Despite constricted space conditions in inner-city areas, the project costs can thus be significantly lowered by an efficient material flow based on a well-planned layout. Normally every development site is liable to various conditions, variation in the flow of construction work, and the local conditions for each construction project are unique. Nonetheless, in real life practices, site layout plans are poorly developed and lack a strategy for minimising repeated handling of materials [6]. In the study carried out by Sadeghpour and Andayesh, numerous techniques are provided for use in jobsite layout development such as knowledge-based approaches (inclusive of heuristics), mathematical programming approaches (including linear and non-linear optimisation), and artificial intelligence approaches (including neural networks, genetic algorithms, and ant or bee-colony and swarm approaches) [7].

In traditional methods, contractors use two-dimensional (2D) drawings, pdf overlays or hand drawings for developing site utilisation plans. Due to the error-prone and poor methods of the traditional 2D-plan-based approach, this tends to lead to non-transparency and the occurrence of mistakes in later processes. There is no digital control of the planning process taking into consideration the specific fringe conditions at the construction site. Today, digital 3D objects can support those responsible for planning and increase planning performance. Furthermore, it is essential to describe the different construction phases and their dynamic fringe circumstances and deduce corresponding suitable measures for the planning of supply and material flow [6]. As each construction project is unique, each construction site layout planning process should be treated as such. Therefore, for every phase of a construction project, sequential consideration of assigned resources in producing site layout plan is essential for the effective, safe, and productive coordination of resources and working environment. Combining the logistics plan of a worksite with a digitalised 3D model forms the grounds for comparing different scenarios and constructs an efficient site layout plan that corresponds with every step of the project.

Building Information Modelling (BIM) is being increasingly used in the construction industry to produce n-D data-rich models. Researchers are attempting to use BIM to promote different aspects of the construction process, including the construction site management and the use of on-site models [8-10]. Deshpande and Whitman suggest that the data from BIM can be leveraged to improve the construction site utilisation planning process and enhance communication among all the stakeholders about the construction execution plan. Also, in their work, they discuss the benefits and contents of the tools that are currently being used by industry professionals and inform readers on how to improve existing tools so that BIM models can be effectively used for site utilisation planning. BIM and VR are the two visualisation technologies that refer to the immersive environment and can visually assess jobsite conditions [11].

Different researchers investigated the applications of VR in the areas of the construction industry such as site layout and planning, rehearsing erection sequences, progress and monitoring of construction processes, evaluation of construction scenarios, inspection and maintenance, and fire safety and access assessment [12, 13]. Even though VR headset technology is broadly appreciated in the technology world by its users, not many studies and practices exist that promotes the use of this equipment for construction jobsite organisation. Therefore, the purpose of this paper is to assess current practices for site layout planning and develop a 3D construction jobsite layout plan model using VR HMD technology, also known as a VR headset. The main objectives of this study are to examine the application of VR headsets in construction jobsite logistics/layout planning, demonstrate and compare traditional 2D site plans, 3D model a site plan on screen and 3D site plan a model in a virtual environment, and explore the potential contribution of VR technology for increasing the productivity and efficiency of jobsite organisation. Within this context, existing jobsite planning applications and the use of VR technology is examined, an on-going building construction project is selected and a 3D model of its traditional 2D jobsite layout plan is conceptualised using Autodesk Revit and Sketchup, a generated 3D model is installed on the VR headset, and a questionnaire survey is conducted with numerous industry professional participants to measure and compare the effectiveness of the highlighted models in the following sections.

2. LITERATURE REVIEW

This section provides an overview of the construction jobsite organisation and VR applications in the construction environment. This information will guide the readers to comprehend the applications of VR in construction jobsites.

2.1. Construction Jobsite Organisation

Among the stakeholders in construction projects, it is usually the project manager's responsibility to prepare a site utilisation plan. These plans are normally developed with the knowledge, understanding, expertise, and intuition of the appointed project manager/planner [14]. Planning of construction jobsite utilisation is a significant task that requires a critical review of alternatives so that optimal decisions can be made about where to locate temporary facilities and how to coordinate people, materials, and plants within the work environment [15, 16]. If a site is not effectively and accurately planned by the project management team, problems that lead to time and cost overruns become inevitable. In traditional practice, the development of a competent site utilisation plan is often stricken by project management tasks such as scheduling, construction method in preference, materials, plant, and labour planning [14]. In spite of the vitality of site utilisation planning, it is often done speedily or occasionally disregarded. The results of such acts are reflected in the day-to-day operations of a project, making management of site operations complicated [17].

Poor site layout planning contributes to inappropriate storage within the worksite that can result in materials and product damage, poor siting of the plant, poor siting of temporary facilities, inadequate space provision, unsatisfactory access, security and safety issues, poor wayfinding, demoralised workers, delays, and increased costs. According to Deshpande and Whitman [15] and Kamat and Martinez [18], early decisions made to develop a good site utilisation plan can have a significant impact on the effectiveness of site operations. This study states that ideal site utilisation plans should be developed in a certain way so that the involvement of labour with the movement of materials is minimised. Hence, they can mainly focus on performing necessary construction activities. According to Mincks and Johnston [19], a site utilisation plan should include clearly designated areas on the jobsite for material delivery, material storage, temporary offices and facilities, jobsite access concerning the movement of material within the jobsite and worker transportation, temporary offices, storage facilities, dry shacks, sanitary facilities, temporary water, power, and heat, and jobsite security. Jobsites that are clean and well-organised provide a working environment that has a positive impact on workforce morale and, in turn, results in high production during each work shift. There are various approaches to developing the job-site layout plan and increasing site optimisation. Each approach has to take into consideration the factors and variables that will lead to jobsite productivity as well as successful contract delivery [1, 7].

2.2. Virtual Reality in Jobsite Organisation

VR replaces the real world with a computer-generated environment and has the ability to model future and current realities. With the rapid developments in VR technology over the past few years, the adoption of this technology by construction industry professionals can contribute significantly to the digitisation of construction sites hence, effectively coordinating the resources allocated for each specific project. VR technology can be used as an important tool for jobsite organisation, which comprises a substantial number of tasks spanning from planning to the execution stage. Presently, some construction firms have initiated the use of VR to plan, manage, and construct their projects. Alternatively, VR technology is used as part of a construction firm's marketing strategy as this technology broadens the vision of potential end-users about what output to expect after the completion of the project. Due to the complexity and the number of factors involved, jobsite developers have adopted novel technologies to improve the efficiency of their tasks. VR technology provides an effective means of verifying site operations as well as site logistics. It allows superintendents to step into the future of their project and understand their requirements and problems long before they occur. Kizil and Joy [20] state that the use of the high-quality 3D environment and dynamic simulation combine to form a uniquely engaging experience. Adoption of VR technology by project management teams enables the visualisation of entire construction jobsite arrangements at different stages thus, providing the workers with a more illustrative site plan by letting them visually see the space utilisation. Jobsite organisation, including traffic routines for equipment, safe working environment, locations for material stock, and manpower availability can also be incorporated into the VR module to be checked visually as part of the logistics plan [21]. Froehlich and Azhar [13] investigated the applications of 3D VR headsets in construction safety and jobsite management. Ebner et al. [6] developed a practice-oriented planning instrument that allows the determination of layout and the equipment for the construction site for the entire running time of the project and it is only changed if needed. Paradoxically, VR technologies are still lagging behind the visions that people have for their use. However, VR has already demonstrated its capacity to change the ways we design, make decisions about, and produce built environments.

3. METHODOLOGY

The focus of this paper is to assess the current understanding about site layout planning through a collective and critical review of literature while building the rationale for why and how VR technology should be considered by decision-makers for production of effective jobsite organisation plans. Therefore, the authors of this article initially performed a comprehensive literature review to investigate the views, perspectives, and practices of site

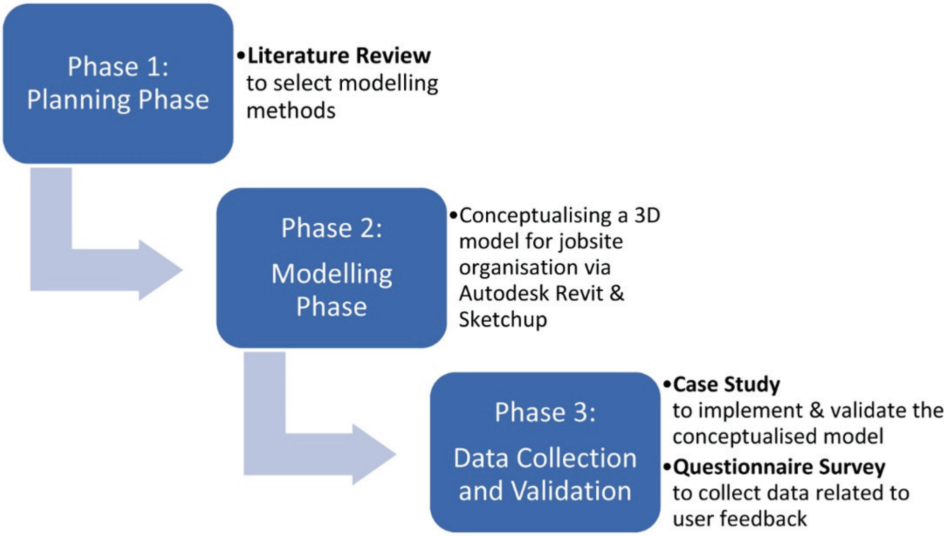


Figure 1 - Sequential Phases of the Adopted Research Methodology

layout planning so that the integrability of VR technology for jobsite organisation and related aspects and challenges can be examined. Figure 1 illustrates the sequential phases of the adopted research methodology.

3.1. Phase 1: Planning

The primary purpose of the planning phase is to identify the existing applications of VR technology in the construction industry. Moreover, examining current practices, aspects, and methods and interpreting them in a specific format so that a 3D model can be established and used for jobsite organisation is also a part of the planning process. At the end of comprehensive literature review, it is concluded that site access, crane placement location, trash waste location, locations of materials, stockpile of excavation, site office, equipment's location, and site orientation are the potential concerns of site layout planning process that can be integrated with VR application for more effective jobsite organisation [2, 3, 16, 18]. As the last step of the planning process, the authors have proposed a conceptual workflow (See Figure 2) that needs to be executed to develop an effective 3D model for application in jobsite organisation.

Figure 2 - Proposed Workflow of the 3D Model for VR Experience

3.2. Phase 2: Modelling

As a part of the modelling process, the initial step was to select a specific construction site where the parameters mentioned above in the planning phase for the jobsite layout planning process were incorporated so that a 3D model could be established accordingly for implementation purposes.

Figure 3 - Location of the Selected Construction Site

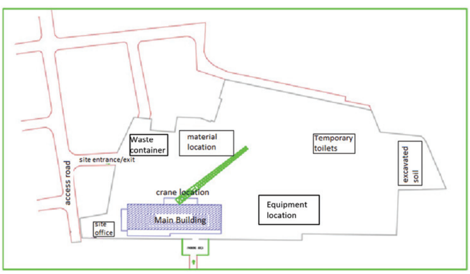
Figure 4 - View of the Main Access Point

For that purpose, the jobsite of a university building construction in the heart of Famagusta (located in Cyprus) city centre is selected to measure the effectiveness of VR technology for this specific construction site's layout organisation. The budget for the selected project is approximately \$4.5 million, and it is nearly 9,000m². The project is planned to be completed on a tight schedule, 11 months after its start date on 20/10/2016. The location of the site is given in Figure 3 and Figure 4 illustrates the site access point from the main road.

After selection, two different scenarios for the construction jobsite organisation of a university building project are proposed for developing the 3D models. With the help of Autodesk Revit 2018, SketchUp 2018, and Lumion 8 software, two different 3D models for the structural phase of the building construction are generated, so that the practicality and efficiency of a VR box headset for developing a precise jobsite layout organisation model are tested and verified. Subsequently, modelling of the main building's structural frame is performed by Autodesk Revit. A generated model is then saved in AutoCAD file format and exported to SketchUp, which provides a wide range of objects from its warehouse. After placing the necessary construction equipment and objects on the jobsite, the 3D model is finalised. Due to its rich object warehouse, SketchUp software contributed significantly to the actualization of the 3D model as it presented the model in a stronger imaginable format. After that, the 3D model is uploaded to Lumion software for the final rendering. A screenshot taken from the Lumion rendering software that illustrates the 3D construction jobsite organisational plan model is depicted in Figure 5. Using the different software packages enabled us to include more detailed features in our model.

Two different jobsite organisational plan scenarios are built based upon the 3D model outputted by the Lumion software. The first scenario is the creation of a 3D model for the actual jobsite organisation plan, which was established by the decision makers of the university building construction. In the second scenario, the actual jobsite organisational plan is modified, and some alterations, such as a change of the site access point, crane placement location, trash waste location, material storage locations, a stockpile of excavation, site office, equipment's location, and site orientation are made. After modelling the different scenarios, VR headsets are used for observing and comparing the scenarios. A 2D site plan and the 3D site plan are modelled based on two different scenarios as shown in Figure 6 and Figure 7. Scenario 1 differs from Scenario 2 based on the site access (entrance and exit), the location of the site office, parking areas, equipment location, crane placement, material locations, trash waste access, and a stockpile of excavated soil. The 3D site model of the two scenarios is used for the VR experience. A 3D site plan model is also created including more features of visualising tools that will enable the observer or viewer to explore the 3D model and also have a real-life experience. In the 3D site plan model, equipment, materials, and other components that are present at the construction jobsite are shown clearly, in detail from the rendered model as shown in Figure 8 and Figure 9.

Figure 5 - 3D Construction Jobsite Organisation Model



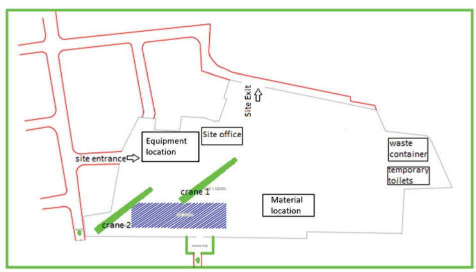


Figure 6 - 2D Site Plan for Scenario 1

Figure 7 - 2D Site Plan for Scenario 2

Figure 8 - 3D Site Plan Model for Scenario 1

Figure 9 - 3D Site Plan Model for Scenario 2

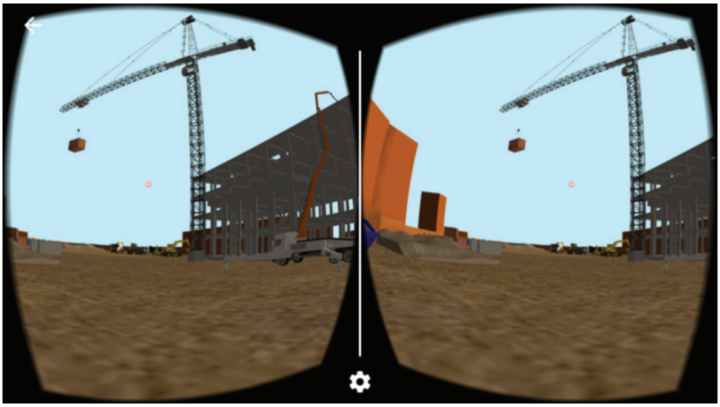


Figure 10 - User's View through the VR Headset for Jobsite Scenario 1

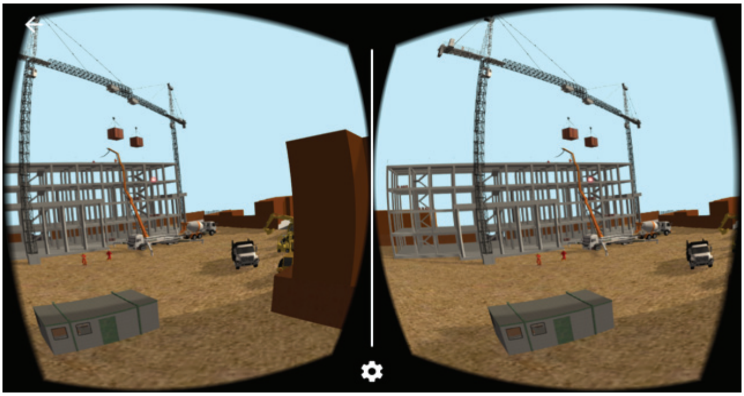


Figure 11 - User's View through the VR Headset for Jobsite Scenario 2

The two 3D site plan scenarios are used for the immersive environments that provide the VR experience for the user. A 360-degree view of the virtual world is experienced through the VR box, through which the user can engage in observing the simulated construction jobsite model. Figure 10 depicts the user's view for jobsite Scenario 1, and Figure 11 shows the user's view for jobsite Scenario 2.

3.3. Phase 3: Data Collection and Validation

In order to measure the reliability of the proposed 3D modelling and integrability of the VR headset technology in designing jobsite organisation plans purposive sampling method was used. This method is described as a type of sampling method in which researcher determines the target population through literature review [22]. In addition to purposive sampling method, through personal communication, experience and observations within the construction sector, 102 industry professionals who are currently using methods similar to proposed framework or whom have the potential to implement this framework are identified. As given in Table 1, field of expertise of the selected participants are composed of engineers, architects, project managers and supervisors. Having participants who practice at different phases of the construction such as planning, executing and controlling for assessment of the proposed framework is believed to enrich this study. Job positions of the participants by the time this questionnaire was conducted is presented in Table 1.

Subsequently, participants are asked to monitor, assess, compare, and rate the effectiveness of the aforementioned university construction site layout plans established in 2D, 3D and 3D in VR with respect to how they perceive and would asess the process. Their ratings are performed based on a Likert scale of 1 (highly ineffective) to 5 (highly effective). The overall ratings of the participants are interpreted to be categorised as high if ≥ 80%, medium if between 60–79%, and low if ≤ 59%.

Table 1 - Demographic Structure of the Participants

Position	Field of Specialisation	Average Years of Experience	Number of Participants
Assistant Site Supervisor	Construction Management	4	8
Revit-BIM specialist	Design Management	5	7
Assistant Site Engineer	Construction Management	3	11
Site Engineer	Construction Management	5	12
Site Supervisor	Construction Management	4	10
Architect	Architectural Design	5	11
Design Manager	Design Management	6	10
Construction Manager	Construction Management	7	11
Structural Site Engineer	Structural Engineering	6	10
CAD specialist	Architectural Design	5	12

4. ANALYSIS AND DISCUSSIONS OF THE RESULTS

In this section, according to the views and opinions of the participants, their ratings on the effectiveness of the proposed site layout planning models (2D, 3D, and 3D in VR) are presented and discussed. Table 2 provides the opinions of the participants for the 2D site layout plan of the university building construction. For both scenarios, ease of use and perceived usefulness of the traditional 2D site layout plan, as well as comprehending how works and related resources will be coordinated within the site, gained the highest average rating from the participants, at 4.1 out of 5. On the other hand, the majority of the participants accept that the 2D site plan (1.9/5) developed for scenario 2 was insufficient, as they believe it might not be possible to detect potential collision between the two tower cranes used on site. Participants have a different opinion concerning the adequacy of the 2D site layout plan for improving the accuracy of jobsite organisation. Some of participants have a moderate response regarding whether 2D site plans provide adequate details of the site planning such as locations of temporary facilities, equipment and materials, as half of the participants have given a rating of 3 for this statement.

Table 2 - Results for Demonstration of Scenario 1 and 2 on 2D Site Layout Plan

	Questions	**Participants Rating (Frequency)**					**Average Rating**
		5	**4**	**3**	**2**	**1**	
1	How easy is the 2D site plan to understand?	39	42	12	9	0	**4.1**
2	How effectively does the 2D site plan facilitate the evaluation of two different site plan scenarios?	14	18	41	29	0	**3.2**
3	How effectively does the 2D site plan indicate variations of proposed jobsite organisation scenarios (1 and 2)?	28	41	33	0	0	**3.9**
4	How efficient is the 2D site plan regarding providing a vision of the expected jobsite organisation in reality?	4	10	15	47	26	**2.2**
5	Is the 2D site plan developed for scenario 2 effective in terms of detecting a potential collision of the two tower cranes?	0	0	27	39	36	**1.9**
6	How well does the 2D site plan show the distance between the tower cranes and the location of the material?	0	35	43	15	9	**3.0**
7	To what extent can the 2D site plan help in improving site layout accuracy?	3	12	44	43	0	**2.7**
8	How well do the 2D site plans show the volume and amount of spaces on site?	5	21	47	29	0	**3.0**
9	To what extent do the 2D site plans provide decision makers with the location details of temporary facilities, equipment, materials, and workflow within the site?	4	18	51	23	6	**2.9**
10	How would you rate the overall effectiveness of a 2D site plan for the proposed jobsite organisation scenarios?	9	22	43	28	0	**3.1**

The participants' ratings on the effectiveness of a 3D site layout plan for both scenarios are presented in Table 3. According to the respondents, the difficulty level of comprehending a 3D site layout plan is above moderate, with an overall rating of 3.9. Moreover, participants believe that site plans developed on a 3D platform contribute greatly to obtaining an effective jobsite organisation plan. When overall ratings of the 2D and 3D site layout plans are compared, participants perceive 3D plans as a more effective instrument since the average rating is 3.9 for the 3D plans, whereas it is 3.0 for the 2D plans.

Table 3 - Results for Demonstration of Scenario 1 and 2 on 3D Site Layout Plan

	Questions	**Participant Rating (Frequency)**					**Average Rating**
		5	**4**	**3**	**2**	**1**	
1	How easy is the 3D site plan to understand?	32	41	19	10	0	**3.9**
2	How effectively does the 3D site plan facilitate the evaluation of two different site plan scenarios?	30	47	25	0	0	**4.0**
3	How effectively does the 3D site plan indicate variations of proposed jobsite organisation scenarios (1 and 2)?	21	52	26	3	0	**3.9**
4	How efficient is the 3D site plan in terms of providing a vision of the expected jobsite organisation in reality?	18	49	30	5	0	**3.8**
5	Is the 3D site plan developed for scenario 2 effective in terms of detecting a potential collision of the two tower cranes?	19	38	43	2	0	**3.7**
6	How well does the 3D site plan show the distance between the tower cranes and the location of the material?	29	55	18	0	0	**4.1**
7	To what extent can the 3D site plan help in improving site layout accuracy?	22	51	20	9	0	**3.8**
8	How well do the 3D site plans show the volume and amount of spaces on site?	17	44	37	4	0	**3.7**
9	To what extent do the 3D site plans provide decision makers with the location details of temporary facilities, equipment, materials, and workflow within the site?	32	49	21	0	0	**4.1**
10	How would you rate the overall effectiveness of a 3D site plan for the proposed jobsite organisation scenarios?	36	59	3	4	0	**4.2**

Results obtained from the 3D jobsite model testing using the VR headset are given in Table 4. First, participants rated the ease of use of the VR system (VR headset) as 2.9. This may be chalked up to the fact that evaluators who have engaged in testing procedures of the 3D model by adopting VR headsets are novice users of this instrument and, due to the complexity of adjusting the headset lenses sideways to provide a clear view for the users to avoid a double image effect and a blurry view through the headset, they may have faced difficulties. One of the participants commented further and stated that the assessment of the system is a bit slower, but it enables anticipation of the reality of the site conditions and restrictions. With a 4.4 mean rating value, participants indicate their belief for the high potential of VR technology for checking problems in the construction jobsite that may arise due to poor jobsite organisation. For Scenario 2, participants are provided with a set of screenshots obtained by the VR headset and asked to observe a collision between two cranes. The materials' stacked location is rated as 3.9 concerning the competency of the VR headset in assisting the visualisation of the distance between the crane positions. This rating proves that most of the participants can visualise the materials in relation to the crane placement location on site. The effectiveness of VR in evaluating the two different jobsite scenarios showed a high rating of 4.2.

VR technology can be used in the evaluation of different construction scenarios. This finding proved that VR technology could be used in different construction site scenario evaluation. Overall, the results indicate that the participants are satisfied with the effectiveness of the 3D model showing the narrowness of jobsite access. The VR system enables users to be immersed in the virtual construction jobsite environment and to view the environment in a 360˚ platform; hence, users are able to observe a realistic experience of reality. It is concluded that most of the participants agreed that VR provides users with real-life-like experience. As a part of this study, most of the evaluators observed the distance between the temporary facilities and the working area through the VR headset. The participants' responses in Table 4 indicate that evaluators are satisfied with the level of detail of the 3D CAD jobsite model in VR, and also that the applicability of VR technology in construction jobsite organisation, which includes studying the site layout plan, the evaluation of different scenarios, and collision detection is found to be promising.

Table 4 - Results of the 3D Jobsite Model Testing Using VR Headset

Questions		**Participant Rating (Frequency)**					**Average Rating**
		5	**4**	**3**	**2**	**1**	
1	How easy is the VR system (VR headset) to use?	9	17	34	39	3	**2.9**
2	How effectively does the VR headset help in checking and verifying problems in jobsite organisation (e.g. a collision between two cranes)?	44	52	6	0	0	**4.4**
3	How well does the VR headset help in identifying that the materials stacked location is far from the second crane in Scenario 2?	29	35	38	0	0	**3.9**
4	To what level of extent does a VR headset provide a better understanding of what to expect on site?	31	44	23	4	0	**4.0**
5	How effectively does the VR technology facilitate the evaluation of two different site plan scenarios?	38	51	13	0	0	**4.2**
6	How well does the VR system help in identifying that the site access (entrance/exit) is too narrow for two trucks to pass at a time?	42	53	7	0	0	**4.3**
7	How well does the VR system help in checking the crane placement location in both scenarios?	35	46	19	2	0	**4.1**

Table 4 - Results of the 3D Jobsite Model Testing Using VR Headset (continue)

Questions		Participant Rating (Frequency)					Average Rating
		5	**4**	**3**	**2**	**1**	
8	To what level of extent can the VR technology help in improving site layout accuracy?	27	39	29	7	0	**3.8**
9	To what level does a VR headset enable the user to feel a life-like experience?	43	50	9	0	0	**4.3**
10	How effectively does the VR headset help in showing that temporary toilets are located far from the working areas?	28	43	27	4	0	**3.9**
11	What level of possible advantages does the VR headset offer in jobsite orientation?	31	47	24	0	0	**4.0**
12	What extent is the level of jobsite organisational plan detailed through the VR headset?	29	58	12	3	0	**4.1**
13	To what extent do you think the VR headset is beneficial when determining construction jobsite organisation?	36	55	11	0	0	**4.2**
14	How well does the 3D VR jobsite model show the volume and amount of spaces on site?	30	49	21	2	0	**4.0**
15	What is the overall rating of the VR technology (VR headset)?	45	54	3	0	0	**4.4**

4.1. Comparison of the Three Construction Jobsite Models

In this study, for the two aforementioned scenarios, three different jobsite layout models are developed, and the participants (comprised of construction industry professionals) are asked for their opinions on the effectiveness of the models via questionnaire survey so a comparison of the models can be progressed. As Figure 12 illustrates, the participants' responses are aggregated for relative frequency.

As depicted in Figure 12, for the questions related to the comparison of the three proposed models, the relative frequency of the 2D modelling is ranked as the least advantageous option. Regarding the detection of the distance between cranes and the location of materials, a collision between cranes on site and evaluation of jobsite scenarios, the participants' preference for the use of 3D models is significantly greater than for the option of using a 2D model. On the other hand, between the ordinary 3D model and 3D model by use of VR technology, the participants' predisposition towards the use of 3D models in VR is slightly higher. Moreover, respondents believe that the 2D model is easier to use for jobsite organisation layout and 3D-VR technology is the most complex option. This necessitates the need for training the novice VR-technology users to mitigate the use of less effective but habitual use of 2D models.

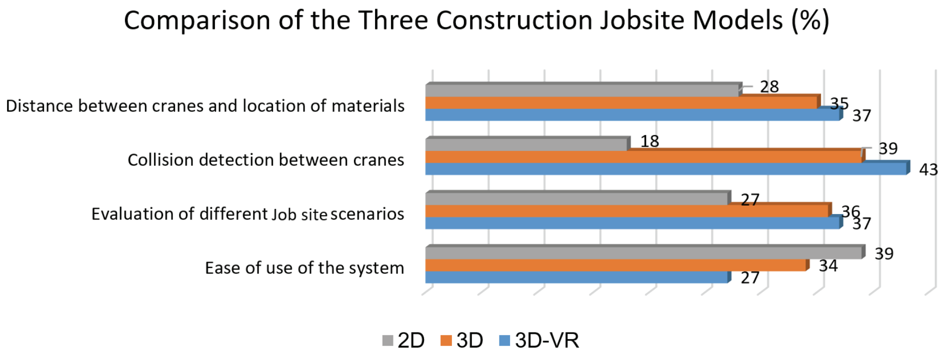


Figure 12 - Comparison of the Three Construction Jobsite Models by Industry Professionals

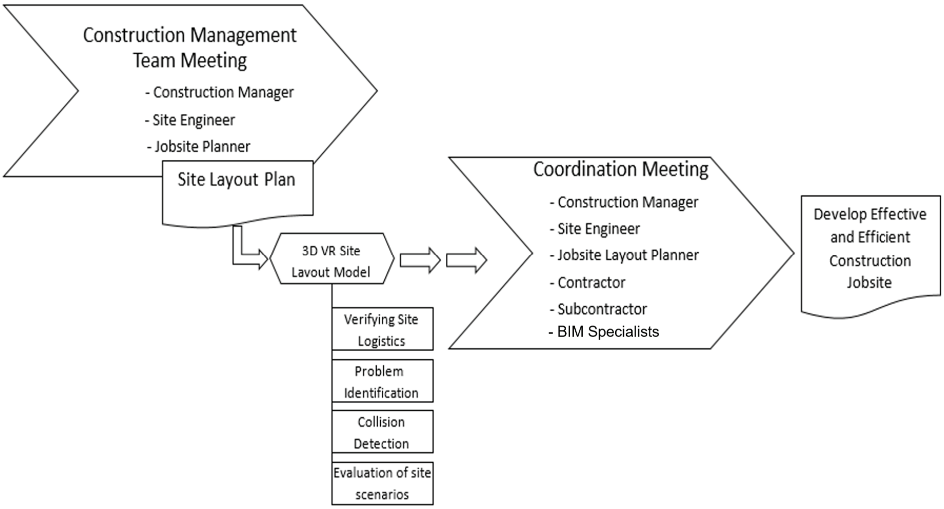


Figure 13 - Proposed Process Flow for Managerial Applications of VR Technology in Construction Jobsite Organisation

5. CONCLUSION AND FUTURE WORK

Construction managers have the key responsibility to manage and allocate individuals for appropriate roles in construction projects. Construction managers (CMs) hold the authority to hire and organise individuals such as site engineers, plant managers, foremen, and labour crews and it is the responsibility of CMs to assign duties to these individuals. Moreover, creating an effective work environment by producing a competent jobsite layout plan is among the critical duties of planners and site engineers. For that reason, whilst allocating individuals with duties, CMs should provide the site engineer and jobsite planners with relevant information on how to improve productivity and safety within the worksite so that it can be reflected on the layout plan. In the proposed process flow given in Figure 13, it is recommended that the final 3D CAD site layout plan model developed by the site engineers and site planners is visualised by a focus group led by the CM using immersive or desktop VR technology. By visualising the model, underlying information of the jobsite organisational plan is believed to be perceived. The visualisation of the jobsite layout plan aids to verify site logistics, problem identification, collision detection, and evaluation of site scenarios in a situation where more than one site scenario is generated. On the other hand, it is essential to schedule a meeting for the related parties so that parties involved in the meeting communicate about matters concerning the establishment of each specific jobsite layout plan. By doing so, problems can be mitigated before the establishment of the actual construction site. Hence, an effective work environment can be developed.

This study seeks to explore the applicability of VR technology for construction jobsite organisation. For this purpose, the comparison between the traditional methods used in construction jobsite organisation and 3D site planning via adopting a VR headset is performed. Within this framework, the ability of VR technology for jobsite layout organisation is explored. The findings reached in this study are based on both direct observations of the participants of the 3D VR site plan models (screenshots obtained from the VR headset) and analyses of participants' questionnaire survey results. At the end of the comparison, the authors of this study discovered that traditional 2D methods used in site layout planning are less time-consuming and easier to understand compared to the VR-based 3D site plans. This could be for a range of reasons, such as inexperienced users and the complexity of different headsets. On the other hand, participants' responses indicate that the ability of the site plans in assisting users to comprehend and evaluate different jobsite scenarios as well as collision detection ability is more effective if a 3D site plan in VR technology is adopted instead of the 2D and 3D site models. Also, participants stated that the assessment of the system is a bit slower, but it enables anticipation of the reality of the site conditions and restrictions. It is recommended for construction managers to practice 3D CAD visualisation technology as it is beneficial to determine how the construction jobsite is being organised. As discussed above, it is to the advantage of jobsite planners to make use of the proposed visualisation technology in other to improve the efficiency of the construction jobsite organisation tasks. This visualisation technology can help construction managers and jobsite planners to visualise the arrangement of jobsites, thus providing them with a more detailed illustration of the site plans and enable them to visualise the amount of space available on site. By visually observing the space utilisation, one can develop a better construction site utilisation plan.

The body of results shows the applicability of the framework for a variety of engineering scenarios, problem-solving, and improving the performance of projects especially for site layout planning. Future directions of work include testing different available types of VR technologies at the market for site layout planning with presented framework. We will also investigate the benefits of this framework in different international construction projects to identify strengths and weaknesses to improve the overall performance of that.

Acknowledgments

The authors would like to thank all companies, senior managers, construction managers, and respondents who contributed to or participated in this research.

References

[1] E. P. Small and M. Baqer, "Examination of job-site layout approaches and their impact on construction job-site productivity," *Procedia engineering,* vol. 164, pp. 383-388, 2016.

[2] A. W. Hammad, D. Rey, and A. Akbarnezhad, "A cutting plane algorithm for the site layout planning problem with travel barriers," *Computers & Operations Research,* vol. 82, pp. 36-51, 2017.

[3] S. RazaviAlavi and S. AbouRizk, "Genetic Algorithm–Simulation Framework for Decision Making in Construction Site Layout Planning," *Journal of Construction Engineering and Management,* vol. 143, no. 1, p. 04016084, 2016.

[4] S. Zolfagharian and J. Irizarry, "Current trends in construction site layout planning," in *Construction Research Congress 2014: Construction in a Global Network*, 2014, pp. 1723-1732.

[5] X. Ning, K.-C. Lam, and M. C.-K. Lam, "A decision-making system for construction site layout planning," *Automation in Construction,* vol. 20, no. 4, pp. 459-473, 2011.

[6] A. Ebner, F. Kammergruber, T. Horenburg, and W. Günthner, "Logistics and layout planning of construction equipment on a vr-multi-touch-tablet," in *Proceedings of 12th International Conference on Construction Applications of Virtual Reality*, 2012.

[7] F. Sadeghpour and M. Andayesh, "The constructs of site layout modeling: an overview," *Canadian Journal of Civil Engineering,* vol. 42, no. 3, pp. 199-212, 2015.

[8] S. A. Salehi and İ. Yitmen, "Modeling and analysis of the impact of BIM-based field data capturing technologies on automated construction progress monitoring," *International Journal of Civil Engineering,* vol.16, no 12, pp. 1669-1685, 2018.

[9] S. Alizadehsalehi, I. Yitmen, T. Celik, and D. Arditi, "The effectiveness of an integrated BIM/UAV model in managing safety on construction sites," *International journal of occupational safety and ergonomics,* pp. 1-36, 2018.

[10] S. Alizadehsalehi, M. Asnafi, I. Yitmen, and T. Celik, "UAS-BIM based Real-time Hazard Identification and Safety Monitoring of Construction Projects," in *9th Nordic Conference on Construction Economics and Organization 13-14 June, 2017 at Chalmers University of Technology, Göteborg, SWEDEN*, 2017, vol. 13, p. 22.

[11] K. K. Han and M. Golparvar-Fard, "Potential of big visual data and building information modeling for construction performance analytics: An exploratory study," *Automation in Construction,* vol. 73, pp. 184-198, 2017.

[12] D. S. Lyne, "Development of virtual reality applications for the construction industry using the oculus rift™ head mounted display," in *Proceedings of the 13th International Conference on Construction Applications of Virtual Reality*, 2013.

[13] M. A. Froehlich and S. Azhar, "Investigating virtual reality headset applications in construction," in *Proc., Proceedings of the 52nd ASC International Conference, Provo, UT*, 2016.

[14] H. M. Osman, M. E. Georgy, and M. E. Ibrahim, "A hybrid CAD-based construction site layout planning system using genetic algorithms," *Automation in construction,* vol. 12, no. 6, pp. 749-764, 2003.

[15] A. Deshpande and J. B. Whitman, "Evaluation of the Use of BIM Tools for Construction Site Utilization Planning," in *50th ASC Annual International Conference Proceedings*, 2014.

[16] A. ElNimr, M. Fagiar, and Y. Mohamed, "Two-way integration of 3D visualization and discrete event simulation for modeling mobile crane movement under dynamically changing site layout," *Automation in Construction,* vol. 68, pp. 235-248, 2016.

[17] E. Elbeltagi and T. Hegazy, "A Hybrid AL-Based System for Site Layout Planning in Construction," *Computer-Aided Civil and Infrastructure Engineering,* vol. 16, no. 2, pp. 79-93, 2001.

[18] V. R. Kamat and J. C. Martinez, "Visualizing simulated construction operations in 3D," *Journal of computing in civil engineering,* vol. 15, no. 4, pp. 329-337, 2001.

[19] W. R. Mincks and H. Johnston, *Construction jobsite management*. Cengage Learning, 2010.

[20] M. Kizil and J. Joy, "What can virtual reality do for safety," *University of Queensland, St. Lucia QLD,* 2001.

[21] W. Kymmell, *Building Information Modeling: Planning and Managing Construction Projects with 4D CAD and Simulations (McGraw-Hill Construction Series): Planning and Managing Construction Projects with 4D CAD and Simulations*. McGraw Hill Professional, 2007.

[22] J.R. Fraenkel, N.E. Wallen, and H.H. Hyun, “How to design and evaluate research in education”, New York: McGraw Hill, 2012.

Teknik Dergi, 2020 9851-9868, Paper 570

The Spatial Distribution of Liquefaction Susceptibility by Logistic Regression Model Adapted for Adapazari, Turkey

Ahmet Can MERT[1]
Ersin AREL[2]
Akın ÖNALP[3]

ABSTRACT

A logistic regression model has been developed for evaluation of soil liquefaction by the use of cone penetration test (CPTu, PCPT) on data collected from Adapazarı, Turkey. The model inputs are the clean sand equivalent normalized cone tip resistance ($q_{c1N,cs}$) and cyclic stress ratio corrected for moment magnitude of 7.5 earthquake ($CSR_{M=7.5}$) that was experienced in 1999. Liquefaction probabilities (P_L) are obtained for each district of the city for which CPTu data is available with the proposed logistic regression model. Average liquefaction probabilities of the depth interval 0-6 m and coordinates (Longitude, Latitude) of CPT soundings were plotted to construct a liquefaction probability map by longitude and latitude. In order to show the effect of depth in liquefaction potential, the obtained liquefaction probability contours were reconstructed by dividing 0-6m depth into three narrow sublayers of 0-2m, 2-4m and 4-6m wherein liquefaction was observed during the earthquake. For each depth interval, liquefaction probabilities of the districts are compared with the observed liquefied and non-liquefied sites in the city after 1999 Adapazarı Earthquake.

Keywords: Liquefaction, CPT, logistic regression, probability of liquefaction.

1. INTRODUCTION

Liquefaction has been investigated extensively during the past 50 years. Conditions triggering liquefaction and evaluation of liquefaction potential have been the main issues of research since the first evidence of liquefied areas were reported for Alaska M_w=9.2 and

Note:

- This paper has been received on July 18, 2018 and accepted for publication by the Editorial Board on March 22, 2019.
- Discussions on this paper will be accepted by May 31, 2020.

- https://dx.doi.org/10.18400/tekderg.445322

1 Istanbul Kultur University, Department of Civil Engineering, İstanbul, Turkey - a.mert@iku.edu.tr - https://orcid.org/0000-0002-2483-1330

2 Istanbul Okan University, Faculty of Engineering, İstanbul, Turkey - ersin.arel@okan.edu.tr - https://orcid.org/0000-0001-9424-2198

3 Istanbul Okan University, Faculty of Engineering, İstanbul, Turkey - akin.onalp@okan.edu.tr - https://orcid.org/0000-0003-1906-0153

Niigata, Japan M_w=7.5 during the 1964 quakes. The developed evaluation models use in-situ test results such as the standard penetration test (SPT) and cone penetration test (CPT) in order that the resistance of soil against liquefaction is defined [1, 2, 3, 4, 5, 6, 7, 8]. The dynamic response of the earthquake is defined as cyclic shear stress (τ_{cyc}) shown in Eq. 1:

$$\tau_{cyc} = 0.65 \frac{a_{\max}}{g} \sigma_{0v} r_d \tag{1}$$

The effect of dynamic loading is represented in the developed evaluation models as the ratio of τ_{cyc} and the effective vertical stress, which is called the cyclic stress ratio (*CSR*) defined as [1]:

$$CSR = \frac{\tau_{cyc}}{\sigma'_{0v}} = 0.65 \frac{a_{\max}}{g} \left(\frac{\sigma_{0v}}{\sigma'_{0v}} \right) r_d \tag{2}$$

where a_{max} is the peak ground acceleration; σ_{0v} and $\sigma_0{}'_v$ are the total and effective vertical stresses respectively; r_d is the stress reduction factor. Most evaluation models follow similar procedures to determine the liquefaction potential. The main objective is to find a boundary using the cyclic loading and the resistance of soil against this load, which reasonably separates the liquefied and non-liquefied data as shown on a simple model in Figure 1.

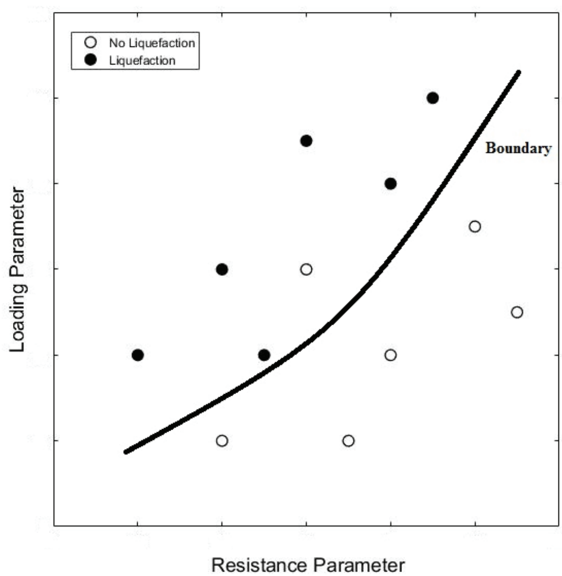


Figure 1 - Simple model for evaluation of liquefaction [*9*]

The deterministic models establish the *Cyclic Resistance Ratio* (*CRR*) and compare it with the existing *CSR* to determine whether liquefaction has materialized. Probabilistic models have also been developed to seek the onset of liquefaction [10, 11, 12]. Logistic regression model is one of the commonly used methods in probabilistic evaluation of liquefaction. Liao et al. [10] applied logistic regression to liquefaction for the first time using corrected SPT blow count ($N_{1,60}$) values. Toprak et al. [11] developed a logistic regression model for liquefaction evaluation with both SPT and CPT data. Since the data from CPTu soundings are more numerous than SPT for a similar borehole depth due to the smaller sampling interval of CPT (2cm/sec whereas one reading per 1.5m in SPT), CPT data have been commonly used in evaluation of liquefaction. Lai et al. [12] proposed a logistic regression model using CSR corrected for 7.5 moment magnitude ($CSR_{M=7.5}$) and normalized cone tip resistance (q_{c1N}). In this study, a logistic regression model has been developed by using $CSR_{M=7.5}$ and sand equivalent normalized cone tip resistance ($q_{c1N,cs}$) in order that the liquefaction susceptibility can be determined as well as the probability of liquefaction. With the developed logistic regression equation, probability of liquefaction for the districts of Adapazari has been obtained by using the CPT soundings carried out after the 1999 earthquake (M_w=7.4). Liquefaction probabilities and available coordinates (Longitude; Latitude) of CPTs have been used to create a liquefaction probability map for the City. The map has been constructed as a two-dimensional planar system which indicates the contours of liquefaction probability. The liquefaction probabilities of each district have been compared with the liquefied and non-liquefied sites in order that the performance of proposed logistic regression equation for liquefaction probability can be tested.

2. LOGISTIC REGRESSION

Logistic regression is used to determine the probability of occurrence of events having binary response. Liquefaction is a convenient topic for the application of the logistic regression because the outputs are either “liquefaction occurs” or “liquefaction does not occur”. The output can be represented as the indicator binary variable Y which consists of “1” for “liquefaction” and “0” for “no liquefaction” in order that the regression can be performed. Y is related to the vector of explanatory variables $\mathbf{X}= [X_1, X_2, \ldots , X_m]^T$. The probability of occurrence is then defined with the help of n number of observations ($\mathbf{X_1}$, Y_1), ($\mathbf{X_2}$, Y_2), …, ($\mathbf{X_n}$, Y_n). In the case of liquefaction, the explanatory variables are determined by considering the representation of cyclic loading effects and the resistance of the soil to these effects. The liquefaction probability P_L with logistic regression is the expected value Y against $\mathbf{X}$ as follows;

$$P_L(\mathbf{X}) = P[Y = 1 \mid \mathbf{X}] = E[Y \mid \mathbf{X}] \tag{3}$$

Since P_L is a probability function, the condition $0 < P_L(\mathbf{X}) < 1$ must be fulfilled, and the regression equation is obtained by solving the logit transformation function Q_L, which has the range $-\infty$ to ∞ defined in Eq. 4 [13]:

$$Q_L = \mathrm{logit}(P_L) = \ln\left(\frac{P_L}{1-P_L}\right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m \tag{4}$$

$\beta_0, \beta_1, \beta_2, \ldots, \beta_m$ are the logistic regression coefficients to be determined with the use of n number of observations. The primary assumptions of logistic regression are that the function Q_L and the explanatory variables X_i are linearly dependent and these variables are normally distributed. The liquefaction probability is then defined in Eq. 5 as:

$$P_L(\mathbf{X}) = \frac{1}{1+\exp(-Q_L)} = \frac{1}{1+\exp(-[\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_m X_m])} \tag{5}$$

Regression coefficients are determined by maximizing the logarithm of likelihood function which consist of the variable vector **X** and the regression coefficients. Likelihood function for liquefaction probability is defined in Eq. 6:

$$L(\mathbf{X};\mathbf{B}) = \prod_{j=1}^{m} [P_L(\mathbf{X})]^Y [1 - P_L(\mathbf{X})]^{(1-Y)} \tag{6}$$

where **B** is the regression coefficient vector.

The goodness-of-fit of the logistic regression has been evaluated by modified likelihood ratio index (MLRI) denoted by ρ^2 as [14]:

$$\rho^2 = 1 - \frac{\ln[L(\hat{\beta})] - \dfrac{(m+1)}{2}}{\ln[L(0)]} \tag{7}$$

where $\ln[L(\hat{\beta})]$ is the maximum value of log-likelihood function, $\ln[L(0)]$ is the log-likelihood value for β_i=0 and m is the number of independent variables. ρ^2 ranges from 0 to 1, and the regression fits well if it converges to 1. Additionally, there is an assumption after Liao et al. [10] that the regression model is well fitted for ρ^2 values greater than 0.4 [12].

3. DETERMINATION OF LOGISTIC REGRESSION VARIABLES

The vector **X** should contain the variables which represent the cyclic loading effect and the soil resistance against this loading. Cone tip resistance (q_c) has been used in this study as the strength parameter. The correction of q_c for overburden is given as the normalized cone tip resistance at 1 atm pressure q_{c1N} [4]:

$$q_{c1N} = C_N \frac{q_c}{P_a} \qquad C_N = \sqrt{\frac{P_a}{\sigma'_{0v}}} \leq 1.70 \tag{8}$$

where C_N is the overburden correction factor and P_a is the atmospheric pressure (95.76 kPa). Another correction to cone tip resistance is made for the fines content (FC). The indicator of the characteristics of soil is the soil behaviour type index (I_c) in CPTu [3] by which the frictional or cohesive behaviour of the soil can be identified. In addition, the effect of fines content is generally represented by I_c for CPTu-based liquefaction evaluations. q_{c1N} has been transformed into sand equivalent normalized cone tip resistance ($q_{c1N,cs}$) by using Eq. 9 [2, 4, 5]:

$$q_{c1N,cs} = K_I q_{c1N} \tag{9}$$

where K_I is the fines content correction factor which is a function of I_c. The cyclic loading effect is generally expressed as *CSR* after the Simplified Procedure in both deterministic and probabilistic studies. Therefore, *CSR* has been used as a loading parameter in the vector **X**. *CSR* has been corrected for the earthquake moment magnitude. The correction has been performed by using

$$CSR_M = \frac{CSR_{M=7.5}}{MSF} \tag{10}$$

where CSR_M and $CSR_{M=7.5}$ are the cyclic resistance ratio for moment magnitude of M and M_w=7.5, respectively. *MSF* is the magnitude scaling factor which may be calculated using the relation proposed by Youd et al. [5] by

$$MSF = \left(\frac{M}{7.5}\right)^{-2.56} \tag{11}$$

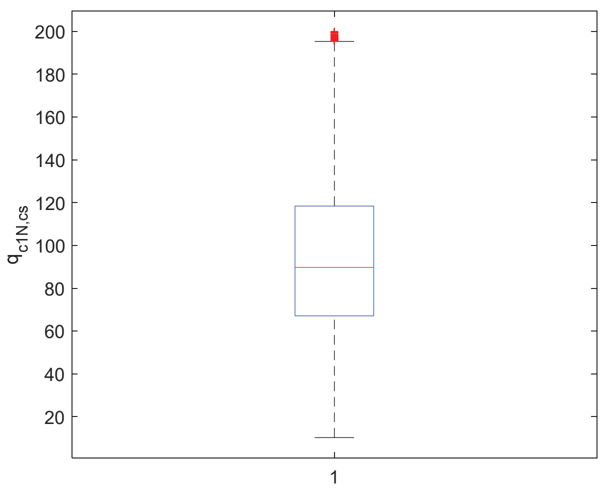


Figure 2 - Box plot for $q_{c1N,cs}$ data

The data used for $q_{c1N,cs}$ has been demonstrated by a box plot given in Figure 2 as to the assumption of logistic regression. The range of $q_{c1N,cs}$ used in the logistic regression model varies from 10 to 200.

In connection with the assumptions involved in logistic regression, the use of square root of $q_{c1N,cs}$ has been preferred [12]. The goodness-of-fit of $q_{c1N,cs}$ and $\sqrt{q_{c1N,cs}}$ to the normal distribution is presented with the probability plots in Figure 3a and b, respectively.

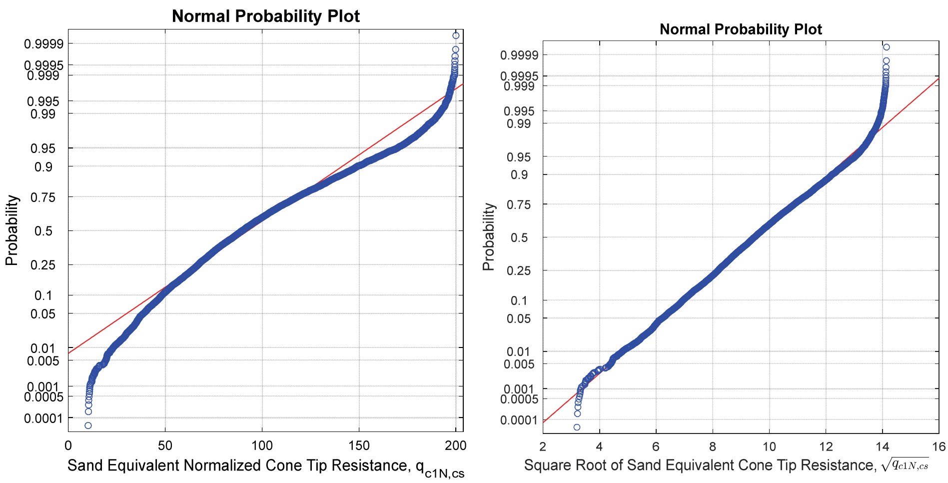


Figure 3 - Probability plot of (a) normalized cone tip resistance and (b) square root of it

Several researchers have expressed the opinion that the use of natural logarithm of *CSR* as explanatory variable increases the goodness-of-fit to the normal distribution [10, 11, 12]. In this study, the variable *CSR* has been used as the form of previous studies ($ln(CSR_{M=7.5})$) in order that the main assumptions of logistic regression are fulfilled.

4. DEVELOPMENT OF LOGISTIC REGRESSION MODEL

Instead of taking the partial derivatives of the logarithm of likelihood function given in Eq. 6 to obtain the regression coefficients, a MATLAB function has been used for this purpose. It is titled "`glmfit` *Generalized Linear Model Regression*" (manual of the MATLAB function is available on https://www.mathworks.com/help/stats/glmfit.html). The data of the explanatory variables have been obtained from the CPT soundings which were carried out in Adapazarı, Turkey after the 1999 [15]. The locations of each CPT soundings and boreholes are depicted in Figure 4.

The total number of q_c data obtained from CPT tests is 9379. The corrections and normalization of q_c have been performed by Eqs. 8 and 9. For the earthquake effect, peak ground acceleration (a_{max}) and moment magnitude (M_w) of 1999 Adapazarı Earthquake records have been used to derive the value of *CSR* using Eq. 1. The corrections for *CSR* have been applied with Eqs. 10 and 11. The indicator binary variable Y has been formed using the "Adapazari Criteria" [16, 17] for liquefaction susceptibility. Fine-grained soil is susceptible to liquefaction whenever the following set of rules are satisfied;

-Liquid Limit (w_L) ≤%33,

-Liquidity Index (I_L) ≥0.9 (instead, use w_n/w_L for NP soils)

-Clay Fraction (D<2μm)<%10,

-Average grain size (D_{50}) >0.02mm

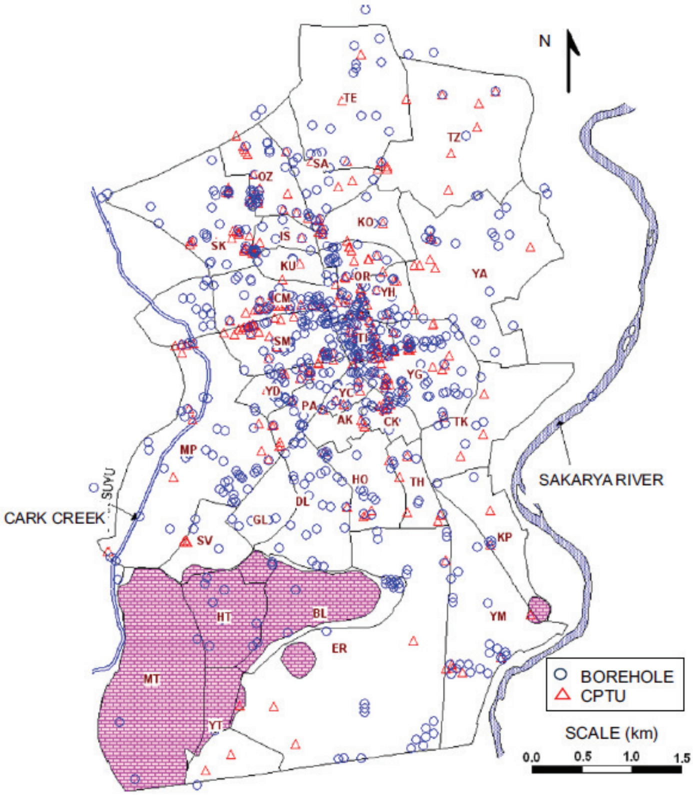


Figure 4 - CPT soundings and borehole locations of the studied area [18]

Adapazari Criteria has been used as indicator binary variable since it has been proven that the liquefaction susceptibility is reflected more accurately than by other studies such as the "Chinese Criteria". Thus, the variable Y consists of 2709 liquefied ($Y=1$) and 6670 non-liquefied ($Y=0$) data. It should be added here that the majority of liquefaction sites have been identified after the earthquake. The database contains physical properties of soil for liquefaction susceptibility and values of explanatory variables from CPT results and earthquake record (Table 1).

Table 1 - Statistical information of the Adapazari database

Variables	**Min. Value**	**Max. Value**	**Range (R)**	**Mean (μ)**	***Coefficient of Variation (COV)***
$q_{c1N,cs}$	10.23	199.98	189.75	94.80	0.41
$CSR_{M=7.5}$	0.25	0.48	0.23	0.38	0.15
a_{max} (g)				0.4*	
M_w				7.4*	
w_L	NP	110	110	32	0.59
I_L	-1.16	8	9.16	0.99	0.75
Clay%	0	70	70	16	0.75
D_{50} *(mm)*	0.001	0.593	0.592	0.042	1.42

* Records of 1999 earthquake (data provided from www.koeri.boun.edu.tr)

The logit function Q_L can be written with the help of the developed regression coefficients as

$$Q_L = 10.4559 - 0.9934\sqrt{q_{c1N,cs}} + 2.476\ln(CSR_{M=7.5}) \tag{12}$$

The relationship of liquefaction probability with logistic regression coefficients is then defined by :

$$P_L = \frac{1}{1+\exp(-[10.4559 - 0.9934\sqrt{q_{c1N,cs}} + 2.476\ln(CSR_{M=7.5})])} \tag{13}$$

The goodness-of-fit of the developed logistic regression model has been found as ρ^2=0.95. The effect of the selection of variable forms to the goodness-of-fit has been presented by comparing MLRIs of the two relations using $\sqrt{q_{c1N,cs}}$ and $q_{c1N,cs}$ as explanatory variables in Table 2.

Table 2 - Comparison of logistic regression equations with different variables

Logit Transformation	β_0	β_1	β_2	ρ^2
$Q_L = \beta_0 + \beta_1\sqrt{q_{c1N,cs}} + \beta_2\ln(CSR_{M=7.5})$	10.4559	-0.9934	2.4760	0.95
$Q_L = \beta_0 + \beta_1 q_{c1N,cs} + \beta_2\ln(CSR_{M=7.5})$	5.9912	-0.0556	2.3479	0.90

Since the logit transformation with explanatory variables $\sqrt{q_{c1N,cs}}$ and $ln(CSR_{M=7.5})$ have greater ρ^2 value according to the comparison given in Table 2, the development of logistic regression model and further studies have been carried out with the probability function of Eq. 13. The boundaries of different liquefaction probabilities have been defined with liquefied and non-liquefied data on $q_{c1N,cs}$ versus $CSR_{M=7.5}$ plane (Figure 5).

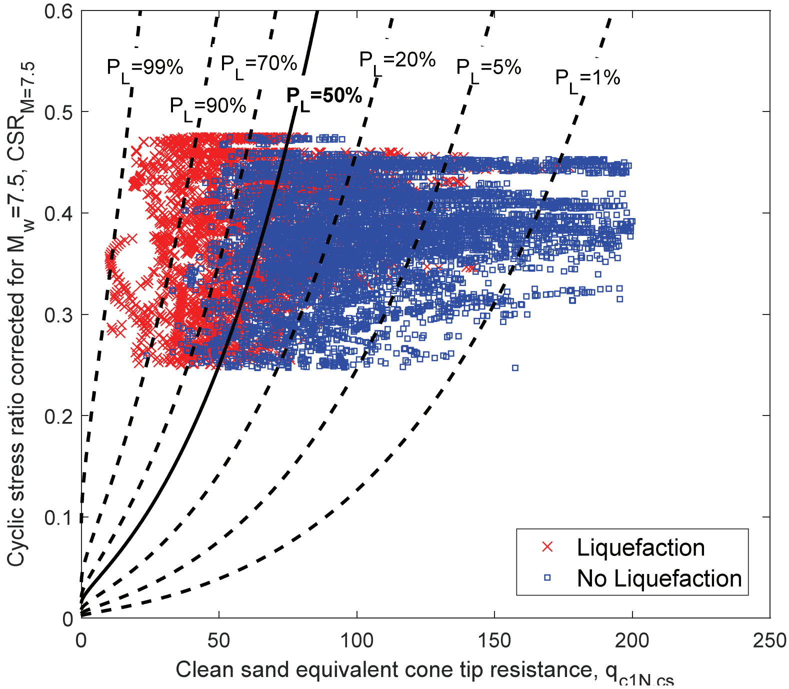


Figure 5 - The developed logistic regression model

In addition to the evaluation of liquefaction potential, the developed logistic regression model provides the possibility of determining the probability of liquefaction with the available inputs of $q_{c1N,cs}$ and $CSR_{M=7.5}$. The boundary of 50% liquefaction probability in the developed logistic regression model is analogous with the boundaries defined in deterministic models since it effectively separates the liquefied and non-liquefied data. Furthermore, the boundaries of 99% and 1% liquefaction probabilities indicate that the zones where liquefaction is most likely and where it has not developed, respectively.

5. DETERMINATION OF LIQUEFACTION PROBABILITIES IN ADAPAZARI

The performance of the developed logistic regression model has been tested by obtaining the liquefaction probabilities for each district of Adapazari. The probabilities have been calculated using the CPT data whose coordinates are made available using the relationship in Eq. 13. The data used in developing the proposed model has been excluded, in order that the soundness of the model can be verified. The liquefaction probabilities for each CPT profile has been plotted on the 2D-planar system of equal liquefaction probability contours for 6m depth along the coordinates (Figure 6a) to secure simplicity of comparison for the liquefied and non-liquefied sites (Figure 6b).

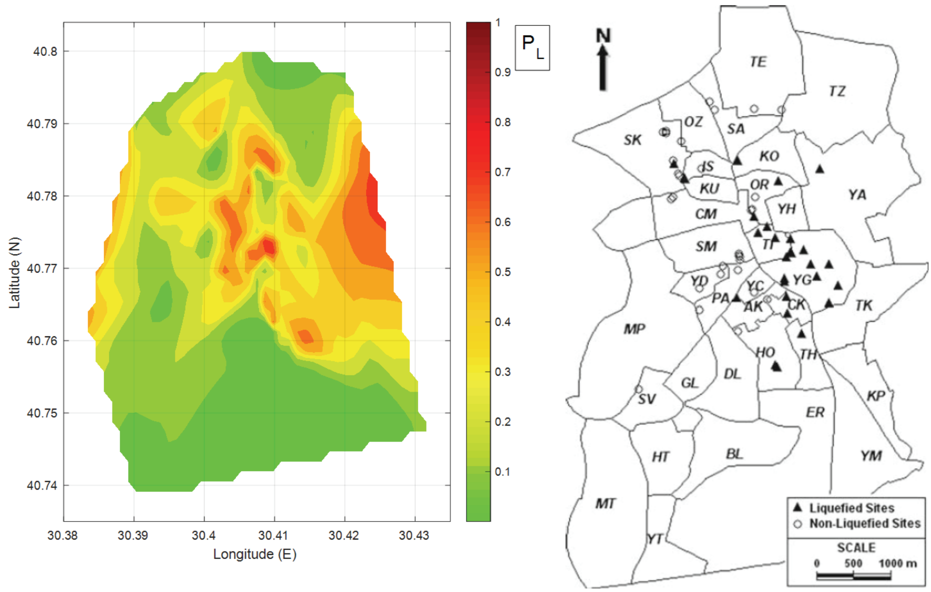


Figure 6 - (a) Probability contour map for 6m depth and (b) comparison with the liquefied sites in the city

In Figure 6b, the liquefied sites are seen to concentrate near the districts TI and YG whose probabilities of liquefaction are the highest at 60-65% as shown in Figure 6a. The non-liquefied sites such as OZ, TE, PA and YD have liquefaction probabilities between 5 and 10%. Although Figure 6a gives analogous information about the liquefaction potential with Figure 6b, it does not reflect the effect of depth on liquefaction because it has been constructed by taking the average of probabilities for the entirety of 6m depth . Therefore, the liquefaction probability contours have been re-established for three sub-intervals of 0-2, 2-4 and 4-6m in order to reflect the effect of depth on liquefaction potential (Figure 8, 9, and 10).

The liquefaction probabilities in the established contours are the highest for shallow depths and diminish with increasing depth. The depth of 0-2 m has the highest liquefaction probabilities whereas 4-6 m depth interval has the lowest, which indicates that the liquefaction has materialized mostly within a few meters from the ground level where Holocene layers as young as 100 years are known to have been sedimented. Figure 7 depicts the typical CPT profile of the liquefied sites at the studying region. In this figure, the resistance increases reasonably at 5 to 6 meters from ground level after which liquefaction cannot be initiated. The soil type becomes clayey after that depth which prevents the triggering of liquefaction.

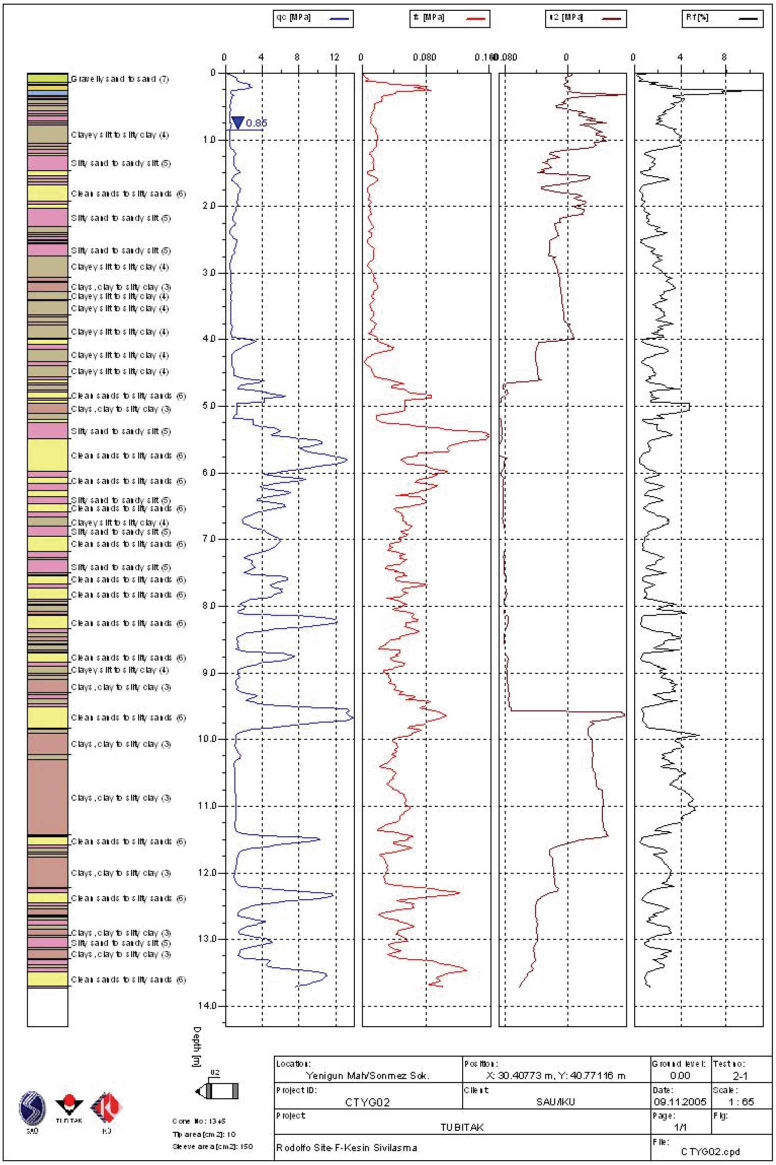


Figure 7 - Typical CPT profile from a liquefied area of the studied region

In Figure 8, the liquefaction probability in the districts where liquefaction was observed to be widespread, have been found to be between 80 and 95%, which confirms that liquefaction occurred mostly near the ground surface. Moreover, the liquefaction probabilities and confirmed liquefied sites match well, confirming that the liquefaction probability estimation by the developed logistic regression model provides reliable results.

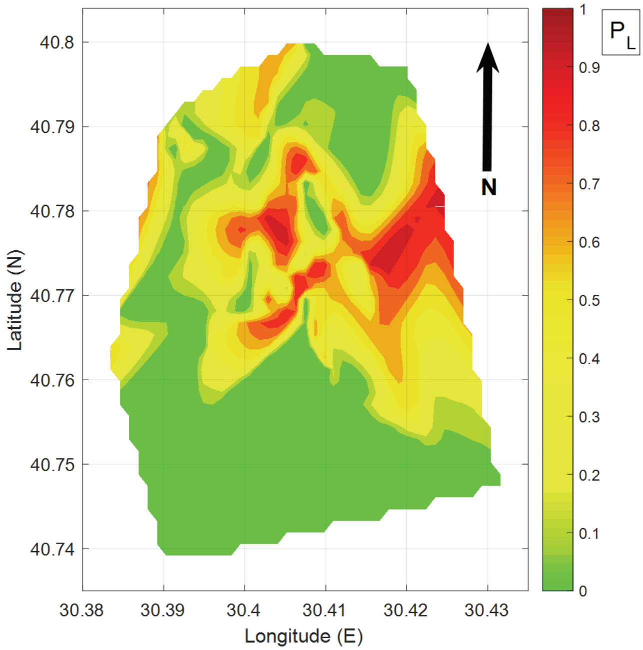


Figure 8 - Equal probability contour map for depth interval 0-2 m

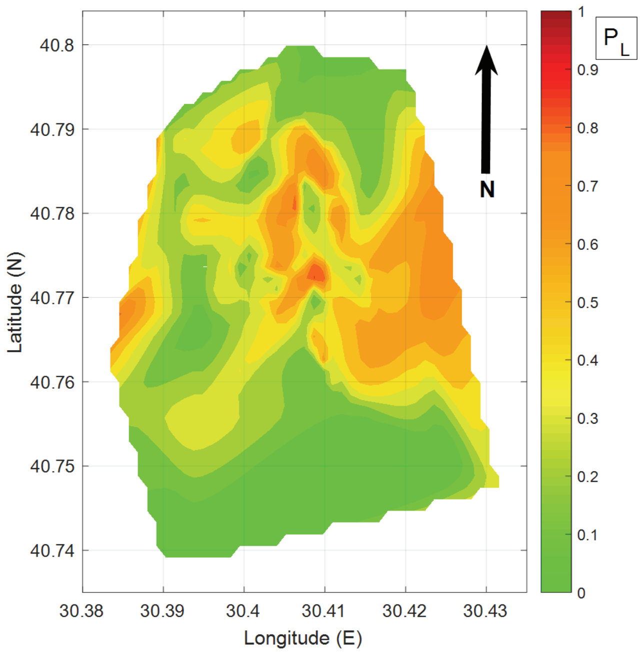


Figure 9 - Equal probability contour map for depth interval 2-4m

The liquefaction probabilities for the coordinates corresponding to the liquefied districts are even lower than for the 0-2 m depth interval in Figure 9 which depicts the sub-layer 2-4m . However, some of the districts have probabilities higher than those obtained for 0-2 m. This may be because either the ground water level (GWL) is within the 2-4 m depth interval or there exist non-liquefiable layers above 2 m. For instance, a location in district TH has 0% liquefaction probability for 0-2 m whereas a probability approximately 50% is indicated for 2-4 m. Nevertheless, the liquefaction probabilities generally do not exceed 70-75% .

Liquefaction probabilities for most of the districts are within 5-30% for 4-6m depth interval. Rare locations in Figure 10 have probabilities of 50-70% for which the liquefaction has occurred along the whole profile. The liquefaction probabilities generally decrease from 0-2 m to 4-6 m depths.

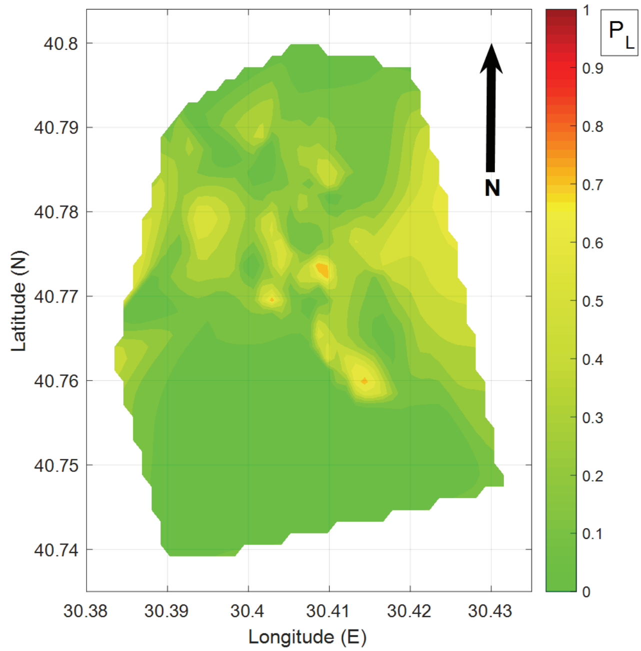


Figure 10 - Equal probability contour map for depth interval 4-6 m

The summary of Figures 8, 9 and 10 are given in tabular form in Table 3. In this table, probabilities of liquefaction have been compared with the available observations of liquefaction for each district. The liquefaction susceptibility is represented as “YES” for liquefaction, “NO” for no liquefaction, and “YES/NO” for the districts where some profiles have liquefied whereas some have not. In the table, cell colours have the same meaning with the colour-scale of the developed probability contours. Red coloured cells indicate liquefied sites, whereas non-liquefied districts have green filled cells. The districts denoted as “YES/NO” as liquefaction condition are assigned yellow coloured cells. In connection with the liquefaction probabilities, those cells with 50% or greater probability are red. Green cells indicate probability of liquefaction below 50% . The cells having both less than and higher than 50% probability are coloured yellow. The colours in Table 3 show that the liquefaction condition in each district and the liquefaction probability for the corresponding district match well. For some of the districts such as KP, "no-liquefaction" is indicated; however, the probabilities remain above 50% for 0-2m depth. The situation emanates from the fact that the soil along the profile may fall within the gray "test" region in according to the Adapazari Criteria. Moreover, some districts such as TH and YC indicate "YES" for liquefaction but probabilities over 50% are encountered for rare depth intervals. This is because liquefaction occurred at a restricted depth of the whole profile.

Table 3 - The liquefaction probabilities for districts of Adapazari

Districts	P_L (%) for depth			Liquefied?
	0-2m	2-4m	4-6m	
AK	84-92	39-63	10-27	YES
CM	6-83	18-57	3-60	YES
ER	-	2-11	0,3-11	NO
HO	74	71	65	YES
IS	-	19	24	NO
KO	93	69	34	YES
KP	54	29	10	NO
KU	20	27	26	NO
MP		62-82	0.2-48	YES
OR	48-93	28-87	13-51	YES/NO
OZ		9-30	14-24	NO
PA		2-8	6-14	NO
SA	7-54	16-58	0-44	YES/NO
SK	18-77	27-72	21-38	YES/NO
SM	5-45	6-25	0-18	NO
SV		30	54	NO
TE		7-34	7-34	NO
TH	28	42-48	39-75	YES
TI	58-93	52-82	13-63	YES
TK	49-78	55-68	6-42	YES/NO
TZ	16	6-25	6-22	NO
YA	59	76	50	YES
YC	27	25-54	0.4-23	YES
YD	24-27	24-47	12-34	NO
YG	60-91	51-82	15-56	YES
YM	4			NO
YT	1.4	1.4	8	NO

6. COMPARISON WITH OTHER LOGISTIC REGRESSION MODEL

The proposed model has been compared with another logistic model in order to confirm that a new model is required for the study of liquefaction susceptibility. The results of a recent study in logistic regression modelling using CPT results [12] have been used for comparison of the findings of this research. Liquefaction probabilities at each district have been calculated using both the developed model in this study and the compared logistic regression equations found in literature. Figure 11 illustrates the discrepancy of liquefaction probabilities of the currently developed and the compared model.

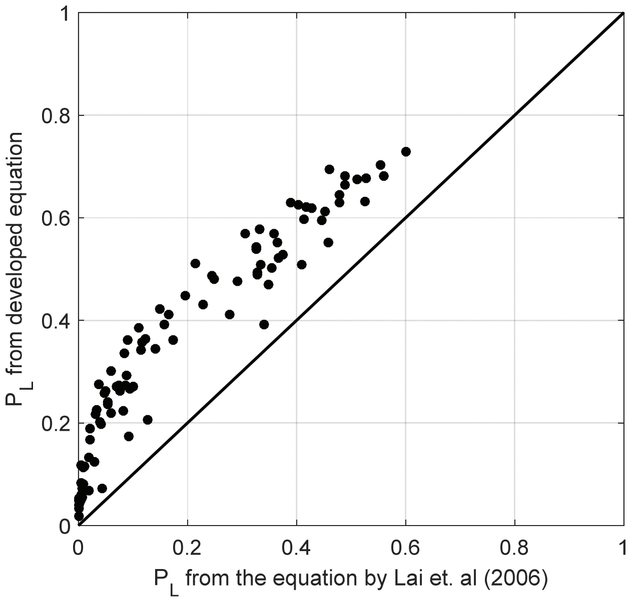


Figure 11 - Comparison of P_L from developed equation (Eq. 13) and the proposed equation by [12]

The equation of the regression model used for comparison gives more conservative values of probabilities than the model developed for this study. It has been found that the use of the previously developed model may result in the underestimation of liquefaction susceptibility in observed liquefaction sites. In contrast, the probabilities derived from this developed model which are above 50% have been verified at the liquefied districts. The maps also illustrate the comparison of average liquefaction probability contours for 6m depth with both developed (Figure 12a) and compared (Figure 12b) models.

The probability contour map from the developed model has indicated values higher than 50% for the liquefied sites whereas the contours from the compared model has less than 50% probability values which do not match with the available observations made at liquefied sites.

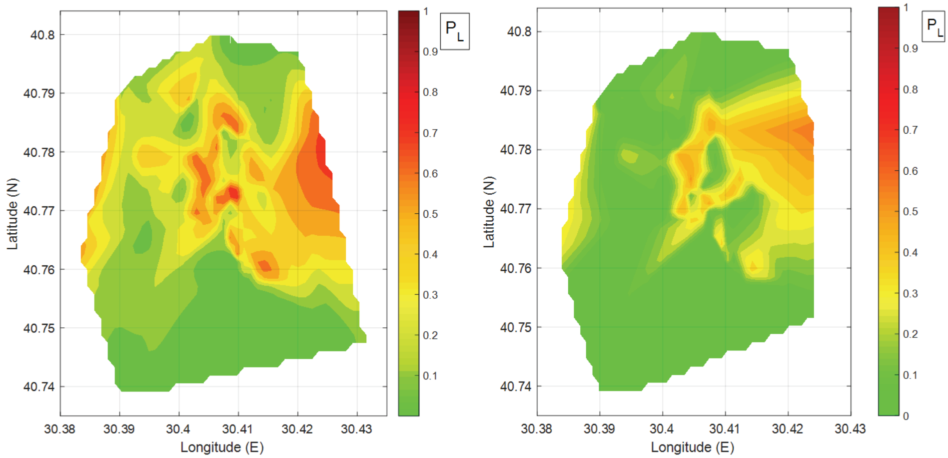


Figure 12 - Equal probability contour maps from (a) the developed model and (b) the compared model

7. CONCLUSION

- A logistic regression model for evaluation of liquefaction susceptibility has been developed with $\sqrt{q_{c1N,cs}}$ and *ln(CSR*$_{M=7.5}$*)* as explanatory variables, derived from 9379 q_c data of 26 CPT soundings performed in most districts of Adapazarı City after the 1999 earthquake as well as a_{max} and M_w values recorded for the earthquake. The developed logistic regression model is as follows:

$$P_L = \frac{1}{1+\exp(-[10.4559-0.9934\sqrt{q_{c1N,cs}}+2.476\ln(CSR_{M=7.5})])}$$

- The goodness-of-fit for the proposed relationship has indicated a high MLRI, at ρ^2=0.95. Average liquefaction probability contours have been constructed for the 0-6 m depth interval for each district in the city.

- The probability contours for sub-layers of the depth of investigation (0-2 m, 2-4 m, and 4-6 m) have been established in order that the change in liquefaction probabilities by depth is reflected.

- It has been determined that the highest liquefaction probabilities are within the top 2m layers. This is due to the finding that the layers through this thickness have been deposited during the floods of the past 500 years.

- The probability of liquefaction has been obtained at 80-95% only for the districts of TI and YG for which the occurrence of liquefaction has been confirmed in situ. Thus, the obtained probabilities with the developed model gives matching results with the observed liquefaction cases.

- For the districts of AK, CM, OR and SK where liquefaction has been triggered in rare cases, the probabilities have been obtained to be both below and above 50%. The remaining districts where the liquefaction has not occurred have the probability less than 50%.
- Among the compared total of 27 districts of the city, the probability of liquefaction has risen above 50% for only 9 districts, which suggests that liquefaction occurred in about 30% of the city during the 1999 Earthquake.
- The necessity for adopting a new liquefaction evaluation approach for Adapazari has been confirmed by comparing the developed model with another logistic regression model from literature. The developed model gives more accurate results than the previous logistic models. In addition, Adapazari Criteria have been used as liquefaction indicator for development of the model which has been proven to represent the properties of ground failure in the region studied.

References

[1] Seed, H. B., Idriss, I. M., Simplified procedure for evaluating soil liquefaction potential, J. Soil Mech. Found. Div., 97, SM9, 1249–1273, 1971.

[2] Seed, H. B., Idriss, I. M., Arango, I., Evaluation of liquefaction potential using field performance data, J. Geotech. Eng., 109, 3, 458–482, 1983.

[3] Robertson, P. K., Campanella, R. G., Liquefaction potential of sand using the CPT, J. Geotech. Eng. 111, 3, 384–403, 1985.

[4] Robertson, P. K., Wride, C. E., Evaluating cyclic liquefaction using the cone penetration, Can. Geotech. J., 35, 442–459, 1998.

[5] Youd, T. L., Idriss, I. M., Andrus, R. D., Arango, I., Castro, G., Christian, J. T., Dobry, R., Finn, W. D. L., Harder, L. F. Jr., Hynes, M. E., Ishihara, K., Koester, J. P., Liao, S. S. C., Marcus, W. F., Martin, G. R., Mitchell, J. K., Moriwaki, Y., Power, M. S., Robertson, P. K., Seed, R. D., Stokoe, K. H., Liquefaction resistance of soils: Summary report from the 1996 NCEER and 1998 NCEER/NSF Workshops on Evaluation of Liquefaction Resistance of Soils, J. Geotech. Geoenviron. Eng, 127, 10, 817–833, 2001.

[6] Boulanger, R. W., Idriss, I. M., State normalization of penetration resistance and the effect of overburden stress on liquefaction resistance, Proc. 11th SDEE and 3rd ICEGE Conf. Berkeley, California, 2004.

[7] Boulanger, R. W., Idriss, I. M., Liquefaction susceptibility criteria for silts and clays, Journal of Geotechnical and Geoenvironmental Engineering, 132, 11, 1413-1426, 2006.

[8] Boulanger, R.W., Idriss, I. M., CPT and SPT based liquefaction triggering procedures, Rep. No. UCD/CGM-14/01, Univ. Of California, Davis, CA, 2014.

[9] Kramer, S. L., Geotechnical Earthquake Engineering, New Jersey, Prentice Hall, 1996.

[10] Liao, S.S.C., Veneziano, D., Whitman, R. V., Regression models for evaluating liquefaction probability, J. Geotech. Eng, 114, 4, 389–411, 1988.

[11] Toprak, S., Holzer, T. L., Bennett, M. J., Tinsley, J. C. I., CPT- and SPT-based probabilistic assessment of liquefaction potential, Proc., 7th U.S.–Japan Workshop on Earthquake Resistant Design of Lifeline Facilities and Countermeasures against Soil Liquefaction. MCEER, Seattle, 1999.

[12] Lai, S. Y., Chang, W. J., Lin, P. S., Logistic Regression Model for Evaluating Soil Liquefaction Probability Using CPT Data, Journal of Geotechnical and Geoenvironmental Engineering, 132, 6, 694-704, 2006.

[13] Cox, D. R., The Analysis of Binary Data, London, Chapman and Hall, 1970.

[14] Horowitz, J. L., Evaluation of usefulness of two standard goodness-of-fit indicators for comparing non-nested random utility model, Advances in trip generation, Transp. Res. Rec, 874, 19–25, 1982.

[15] Önalp, A., Arel, E., Bol, E., Özocak, A., Sert, S., Application of Dissipation Method in Cone Penetration Test (CPT) for Determination of Liquefaction Potential (In Turkish), 104M387, The Science Foundation (TUBITAK), Adapazari/ Sakarya, December 2007.

[16] Önalp, A., Arel, E., The liquefaction potential of silts-The Adapazari Criteria (in Turkish), ZM9 9th National Congress on Soil Mechanics and Foundation Engineering. Eskisehir, Turkey, 2002.

[17] Bol, E., Önalp, A., Arel, E., Özocak, A., Sert, S., Liquefaction of Silts: The Adapazari Criteria, Bulletin of Earthquake Engineering, 8, 4, 859-873, 2010.

[18] Arel, E., Predicting the spatial distribution of soil profile in Adapazari/Turkey by artificial neural networks using CPT data, Computers and Geosciences, 43, 90-100, 2012.

Teknik Dergi, 2020 9869-9896, Paper 571

Time - Cost Relationships for Superstructure Projects in Turkey

Murat SÖNMEZ[1]
S. Ümit DİKMEN[2]
Rıfat AKBIYIKLI[3]

ABSTRACT

The concept of time-cost relationship in construction projects was first introduced by Bromilow, using the data of the 328 superstructure projects completed in Australia. The aim of this study is to determine the time-cost relationship of superstructure projects in Turkey. Time and cost data of 460 superstructure projects completed between the years of 1999-2018 was used in the study. Data was grouped primarily on the basis of the intended use of buildings (individual buildings, educational buildings, hospitals, industrial buildings and social housing), and then time-cost relationships were separately determined for each group through statistical analysis. In addition, the effects of the parameters such as exchange rates and the number of non-working days on the time-cost relationship were investigated statistically, as well. As a result of this study, highly meaningful time-cost relationships (**R^2 =0,60**) are determined for the construction projects in Turkey, excluding disaster relief projects, which is not part of this study. It has also been determined that while the hospital projects in Turkey have the highest coefficient of determination (**R^2= 0,87**), social housing projects have the lowest (**R^2= 0,54**).

Keywords: Bromilow’s time-cost (BTC) model, building projects, construction management, duration estimation, regression analysis.

Abbreviations

BTC **:** Bromilow's time - cost

C **:** Cost

CBRT: Central Bank of Turkey

ER **:** Exchange Rate

Note:

- This paper has been received on September 03, 2018 and accepted for publication by the Editorial Board on March 18, 2019.
- Discussions on this paper will be accepted by May 31, 2020.

- https://dx.doi.org/10.18400/tekderg.456761

1 Sakarya University, Institute of Natural Sciences, Sakarya, Turkey - murat.sonmez2@ogr.sakarya.edu.tr - https://orcid.org/0000-0003-2328-6817

2 MEF University, Civil Engineering Department, Istanbul, Turkey - dikmens@mef.edu.tr - https://orcid.org/0000-0002-4003-1368

3 Duzce University, Civil Engineering Department, Duzce, Turkey - rifatakbiyikli@duzce.edu.tr - https://orcid.org/0000-0003-1584-9384

LTF	**:** Love et al.'s Time-Floor	**NWD**	**:** Non-Working Days
PPI	**:** Producer Price Index	**PRI**	**:** Private
PUB	**:** Public	**T**	**:** Time
TOKİ	**:** Public Housing Administration		

1. INTRODUCTION

In construction, the prediction of a reasonably accurate time and cost relationship is necessary for a successful planning at all phases of the projects [1]. The idea of investigating the time-cost relationship of construction projects was first proposed by Bromilow in 1969 [2]. Bromilow suggested that there is a strong relationship between the realization times of the projects and their cost in his study on building projects in Australia. This relationship proposed by the researcher is known as Bromilow's time - cost (BTC) model in the literature. This result is undoubtedly a very important guide both for the investors and the contractors in estimating the duration of the project as far as possible [3]. Bromilow has also stated that the time and cost variables of the BTC model can be affected and varied at various levels by a number of factors such as the type of project, the climate conditions of the location where the project is realized, the changes in the exchange rate, and the terms of the contract [2]. Subsequently, the model proposed by Bromilow has been studied by numerous researchers at different parts of the world, especially taking into account the influence of local factors [2, 3]. Although several new approaches such as using Artificial Neural Networks [4], reference group class [5], genetic algorithms [6] etc., proposed for predicting time cost relationship in construction projects, the BTC approach still is the widely accepted approach due to its simplicity.

The aim of this study is to determine the time-cost relationship of superstructure projects in Turkey utilizing the data of the projects completed by public and private sectors between years 1999-2018. In the study, data of these projects are categorized under five different major types, namely hospitals, educational buildings, social housing, individual buildings and industrial building projects. In the study, the effects of the periodical exchange rates and the number of non-working days were also investigated. In this context, the data of the 460 completed projects were analyzed. Consequently, remarkably good results were obtained which can be utilized in the construction sector for project planning in Turkey.

2. RELATED WORK

2.1. Bromilow’s Time - Cost Model

As mentioned above, the idea of a possible relationship between the realization cost and the construction duration of projects was first proposed and researched by Bromilow [2]. In his study in 1969, Bromilow examined the time-cost relationship of 328 superstructure projects in Australia. The researcher mathematically expressed the relationship between the cost and construction time of the projects by the following equation,

$$T = KC^B \tag{1}$$

where, T is the time period between the calendar day that the project starts with the contract signing and ends with the substantial completion of the project. C is the final cost for the project to owner, namely the investor at the end of the project. In other words, it is the total payment received by the contractor. On the other hand, K and B are the coefficients that correlate time and cost. K is a characteristic of the construction time performance (depending on the country), and B is the value that shows the performance of the construction time versus the financial sensitivity. Though Equation (1) is an exponential equation, it can readily be transformed into a linear form as follows,

$$\ln T = \ln K + B \ln C \tag{2}$$

Hence, now the coefficients K and B can easily be calculated by regression analysis. Bromilow, for Australia, determined the K coefficient as 350 and the B coefficient as 0.30 in his study (Figure 1). The researcher assigned the value of C in AU$ 1 million. Based on these coefficients, for example, a project with a cost of AU$ 1 million constructed in Australia in 1969 would require approximately 350 days.

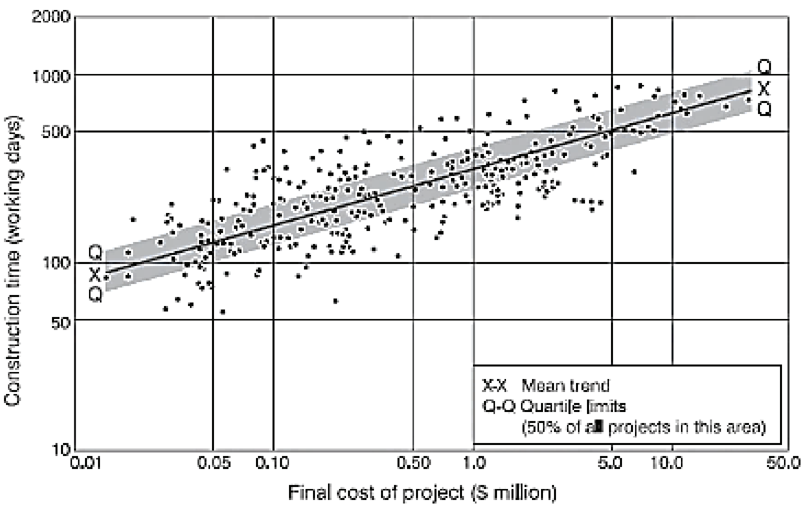


Figure 1 - Bromilow's original data set (1969) [1]

2.2. Related Studies in Other Countries

Numerous researchers, by using data sets from different countries, conducted studies to determine the country specific K and B values of the BTC model. A summary of these studies demonstrating K, B and R^2 values is presented in Table 1. The R^2, a statistical value, indicates the strength of the relationship between the variables, namely cost and time. An $R^2 = 1$ indicates that the time-cost relation is fully explained, whereas if the value is 0, it can be concluded that there is no relationship [8].

Love et.al. (2005) considered the physical characteristics as an important determinant of the construction time and proposed the LTF (Love et al.'s Time-Floor) model derived from the

BTC model for calculating construction time by using the total floor area and number of floors [9]. Ng et al. (2001) found that unit construction periods developed positively over time, when compared to previous findings in Australia, and they attributed this improvement to increased productivity in construction projects over time after Bromilow's study [2, 11]. Kaka and Price (1991) tested the BTC model for public and private sector projects and consequently claimed that the time cost relationship did not vary between the public and private sectors [12].

Table 1 - Equations obtained by other researchers with the Bromilow BTC Model (adapted from Kenley (2003) [40]).

Country	Author	Year	K	B	R^2	Num. Of Projects	P. Type	Currency
Australia	Bromilow [2]	1969	287.0	0.30	-	329	Pub.+Pri.	AUD
	Bromilow [19]	1974	313.0	0.30	-	370	Pub.+Pri.	AUD
	Ireland [9]	1985	219.0	0.47	0.58	25	Pub.+Pri.	AUD
	Yeong [20]	1994	269.0	0.22	-	87	Pub.+Pri.	AUD
	Ng vd. [11]	2001	130.9	0.31	0.58	93	Pub.+Pri.	AUD
USA	Hoffman et.al. [14]	2006	26.8	0.20	0.34	580	Pri.	USD
	Choudhury et.al. [9]	2003	19.0	0.39	0.74	55	Pri.	USD
Bangladesh	Choudhury et.al. [9]	2002	5.0	0.27	0.65	35	Pub.	BDT
	Mizanur et.al. [21]	2014	156.0	0.36	0.65	63	Pub.+Pri.	BDT
Bosnia and Herzegovina	Zujo ve Pusic[22]	2008	70.0	0.52	0.39	29	Pri.	HRK
	Zujo ve Pusic[22]	2008	79.0	0.41	0.51	24	Pri.	HRK
	Petruseva et.al. [23]	2013	0.1	0.55	0.73	75	Pri.	BAM
China	Sun et.al. [24]	2010	42.4	0.29	0.52	72	Pri.	CNY
Indonesia	Sriana et.al.-Bireun[25]	2015	24.1	0.14	0.09	105	Pub.	IDR
	Sriana et.al.-Pidie[25]		19.0	0.15	0.12			IDR
	Sriana et.al.-Utara[25]		12.8	0.26	0.83			IDR
	Sriana et.al.-Selatan[25]		24.5	0.13	0.52			IDR
	Sriana et.al.-Barat[25]		6.68	0.36	0.76			IDR
	Sriana et.al.-Timur[25]		99.3	-0.13	0.05			IDR
	Sriana et.al.-Tengah[25]		15.7	0.21	0.37			IDR
	Sriana et.al.-Banda[25]		53.8	-0.03	0.01			IDR
Ghana	Ameyaw . [26]	2012	3.2	0.38	0.38	62	Pri.	GHS
S.Korea	Le-Long et.al.. [27]	2009	341.0	0.18	0.76	34	Pub.+Pri.	KRW
India	Choudhury[28]	2009	4.4	0.28	0.37	50	Pri.	INR
India	Car ve Radujkovic[29]	2010	58.0	0.50	0.71	107	Pub.	HRK
Hong Kong	Kumaraswamy et.al. [13]	1995	152.0	0.22	0.85	111	Pub.+Pri.	HKD
UK	Kaka ve Price[12]	1991	486.7	0.21	0.68	661	Pub.	GBP

Table 1 - Equations obtained by other researchers with the Bromilow BTC Model (adapted from Kenley (2003) [40]). (Continue)

Country	Author	Year	K	B	R^2	Num. Of Projects	P. Type	Currency
Kuwait	Jarkas[30]	2016	28.8	0.19	0.80	113	Pri.	USD
			7.9	0.28	0.85	74	Pri.	USD
Malaysia	Yeong[20]	1994	518.0	0.35	-	51	Pri.	AUD
	Endut et.al. [16]	2006	328.0	0.25	-	301	Pri.	MYR
	Chan[15]	2001	269.0	0.32	0.41	51	Pub.	MYR
Nigeria	Ojo[18]	2001	27.0	0.13	0.18	-	Pri.	-
	Ogunsemi et.al. [17]	2006	63.0	0.26	0.21	87	Pub.+Pri.	NGN
	Waziri[28]	2014	2.8	0.54	0.54	32	Pub.	NGN
Slovakia	Mačková et.al. [32]	2014	384.0	0.26	0.77	28	Pri.	EURO
Mixed	Özçekiç[33]	2007	137.3	46.05	0.76	209	Pri.	USD
Turkey	Odabaşı[7]	2009	161.7	0.58	0.93	7	Pub.+Pri.	AUD
	Bayram[3]	2016	209.0	0.35	0.26	424/530	Pub.	TRY
Vietnam	Le-Hoai et.al. [34]	2009	98.1	0.34	0.44	77	Pub.	VND
	Le-Hoai et.al. [34]	2009	87.2	0.35	0.38		Pri	VND

Currencies are shown in accordance with ISO 4217 standard [37]. *Pub.= Projects of public sector; Pri.= Projects of private sector; Pub.+Pri.= Projects of public and private sector (Table 1.) Choudhury et al. (2003) used 1000 USD for project costs in their analysis* [

Kumaraswamy et al. (1995) [13] conducted a study by using the BTC model on the 111 projects in Hong Kong. They analyzed these projects in the data set under public and private sub-groups. They confirmed the usability of the model proposed by Bromilow. Hoffman et al. (2007) focused on the Air Force-financed facility projects and indicated that the BTC model represents the data set to a large extent, and they suggested that the BTC model can be used in order to determine time-cost relationship.

Kaka and Price (1991) categorized the projects as public and private sector industrial projects; Kumaraswamy et al. (1995), Chan (2001), Ng et al. (2001), Endut et al. (2006) Ogunsemi and Jagboro (2006) categorized the projects in their studies according to their intended use (residential, commercial educational) [11,12,13,15,16,17]. While, in general, the existence of a time-cost relationship in the construction projects is confirmed by these studies conducted in various countries by using BTC model, some studies also suggested that such a time-cost relationship may be inadequate for some other countries.

Ojo (2001) and Ogunsemi and Jagboro (2006) stated that the BTC model has poor forecasting capability for Nigerian construction projects [17, 18]. The uncertainties in the country, the regional factors and the fluctuations in the economy have been shown as the main reasons for the weak relationship between time and cost. In this respect, the researchers investigated the projects in their data set and proposed a fragmented structure, as an alternative to BTC model, from the breakpoints of the projects and obtained more meaningful results in terms of the studied region.

On the other hand, Ireland (1985) proposed a model to determine the speed of construction (Speed), taking into account physical factors such as gross floor area per day and number of storeys to be constructed [8, 9].

$$\log SPEED = -5.72956 + 2.96889(\log_{10} AREA)^{0.6124} + \frac{2.93390}{Storeys} \tag{3}$$

All the studies carried out with different data sets are presented in Table 1, except Yeong (1994) and Chan (2001)'s studies. These researchers conducted their studies for Malaysia and used the same data set to form the BTC model, but they found different *K* values with the same data set as seen in Table 1 [15, 20]. The fundamental reason for the discrepancy of the *K* coefficients between two studies (Yeong's *K*=518 and Chan's *K*= 269) is that Yeong (1994) transformed the project costs into AUD during his study [20]. On the other hand, as can be seen from Table 1, the *B* coefficients did not show any change when compared to the *K* coefficient. The main result that can be construed from this situation is that while the currency directly affects the coefficient *K* in the model, on the other hand currency does not affect the coefficient *B* in the BTC model. Additionally, it is understood that models can be compared using *B* coefficients, even though the models, created for different countries, use different currencies. Okere (2018) aimed to study time-cost relationship of road projects realized in the Washington State by using the BTC model. The researcher found positive correlation (R^2 = 0,695) between time and cost of the road projects [35]. Adeyemi and Masalila (2016) investigated time-cost relationship of construction projects in Botswana. Researchers found that BTC model can reveal time-cost relationship (R^2= 0,817) of construction projects realized in the Botswana [36].

In conclusion, as can be seen from Table 1, while there are studies that confirm the relationship between time and cost, there are also studies that show low correlation between these two variables. In the literature survey, it was found that the variability of economic conditions, working styles and technical qualifications in the regions where the data set of the studies were collected affect the time and cost relationship significantly. Therefore, taking into account these basic factors that affect the construction projects can provide more accurate results in terms of time-cost relationship.

2.3. Time - Cost Model Studies in Turkey

In Turkey, several studies were conducted. Bayram (2017) investigated the completion period of the public buildings, consisting of health and educational buildings, in Turkey using both Bromilow's time-cost (BTC) and Love's Time - Floor (LTF) models. The researcher concluded that the BTC model is superior to the LTF model, stating that "cost" is a more significant predictor of duration than number of floors and total floor area [6]. Odabaşı (2009) examined the construction projects, consisting of 7 educational buildings, carried out between 2004 and 2007 at the Middle East Technical University Campus and the factors affecting the construction time. Odabaşı suggested that the BTC model can be used in predicting time-cost relationship in construction projects [7].

3. METHODOLOGY

3.1. Data Set

As presented in Table 1, researchers have studied data sets by classifying according to customer types, namely public and private, as well as classifying according to the function of the structure, such as hospital and educational buildings [38]. In this study, during the preparation phase of this work, data of 501 projects carried out in 56 cities in Turkey between 1999 and 2018 has been collected (Table 2).

Table 2 - General Features of Projects

Project Type	Sector (Public (Pub.) / Private (Pri.)	Num. of Projects	Project %	Realization Period of Projects	Shortest Contr. Duration (day)	Longest Contr. Duration (day)	Lowest Contr. Value (Million TL)	Highest Contr. Value (Million TL)
Social Housing	Pub.	130	28.3%	2005 - 2017	203	750	1.28	147.55
Educational Building	Pri.+Pub.	177	38.5%	2007 - 2016	145	685	1.92	103.20
Hospital	Pri.+Pub.	40	8.7%	1999 - 2014	350	1200	0.35	149.50
Individual Building	Pri.	58	12.6%	2011 - 2016	163	773	0.22	120.00
Industrial Building	Pri.	55	11.9%	1999 - 2017	190	650	0.40	79.67
All Projects	Pri.+Pub.	460	100%	1999 - 2017	145	1200	0.22	149.50

The data collected for each project consists of the date of the contract, date and place of delivery of the project, the date of completion of the work according to the contract, the contract value, date of the temporary acceptance, the location of the project, the employer and the contractor firm, the actual payments (cash flows), realization date and cost of the Project. Project data was obtained by interviewing project managers of the Public Housing Administration (Toplu Konut İdaresi-TOKİ) and 65 construction companies operating in the sector. TOKİ projects consist of TOKİ Disaster Housing projects and TOKİ Social Housing projects. However, Disaster Housing projects were excluded from this study because they were built under extraordinary conditions that occurred after the disaster. Therefore, 41 projects were removed from the data set leaving 460 projects to be used in the analysis. The project groups are hospitals (8.70%), industrial buildings (11.96%), individual buildings (12.61%), social housing (28.56%) and educational buildings (38.48%) as tabulated in Table 2. The shortest completion period in the data set is 145 days and the longest is 1200 days. The minimum and maximum contract values are 0.22 and 149.50 Million TL at the time of contract signing. The average contract value of the projects in the dataset was determined to be approximately 20 Million TL and the realized completion cost is calculated to be approximately 21 Million TL. The average completion time of the projects was determined as 422 days. On the other hand, the data set contains both private and public projects. Hence the data set used in this study is a much more comprehensive set than the earlier studies made on Turkish projects.

Projects in the data set were classified in 3 main groups based on their types of use, exchange rates at the time they were built and the non-working days of the cities where they were built. These 3 main groups also have their respective sub-groups as will be explained below.

3.2. Data Analysis

The projects in the data set were realized at different times. Hence, the project costs were affected by the changing economic conditions. Within this framework, the realization cost of the projects should be free from the time effect via indexes and be made comparable on a common basis as recommended in the Bromilow's (1969) BTC model. Therefore, the Equation (4) is used in this study in order to remove time effects from the realization costs.

$$Adjusted\ Real\ Cost = Real\ Cost\left(1 + \frac{I_1 - I_2}{I_2}\right) \tag{4}$$

where, I_1 refers to the index value at the contract date, and I_2 refers to the date at which the project costs are to be examined. Whereas, *Real Cost* shows the realization cost of the project at the time of delivery and the *Adjusted Real Cost* shows the cost of project free of the time effects. In this study, the project costs were adjusted to April 2018 by using the producer price index (PPI) provided at the web site of TUİK.

All the data analysis was performed using Mathworks Matlab version R2017a [39].

3.3. Grouping per Exchange Rates and Number of Non-Working Days

Ireland (1985) emphasized the utility of the BTC model, while underlining the fact that adding variables that measure the time, cost, and quality to the BTC model, would provide more efficient results for projects [9]. Ojo, (2001) and Ogunsemi and Jagboro (2006) showed that the BTC model cannot achieve much success because of the fluctuations and uncertainties in the construction sector and also put forward that alternative mathematical approaches that fit with conditions, could be used for predicting time-cost relationship [17, 18]. Kumaraswamy and Chan (1995) reported that the construction duration was affected by the project location [10]. Finally, Kenley (2003) suggested that the implications and roles of project classifications in prediction models should be addressed by researchers in further works [40].

On the other hand, it is known that the construction sector in Turkey is affected by the exchange rates due to high use of imported materials [41, 42]. Hence, examination of the effect of exchange rates is important. To study their potential effect, exchange rate information was obtained from the website of the Central Bank of Turkey (CBRT). The exchange rates were normalized dividing by the January 2018 rates (Figure 2) [42].

As can be seen from Figure 2, there are 3 different periods in terms of exchange rate trends. Between 1999 and 2003 is the first period when exchange rates displayed a continuous rising trend, on the other hand between 2003 and 2010 is the second period when exchange rates stayed relatively stable. In the third and the last period, between 2010 and 2018, exchange rates had both fluctuating and rising trends. At first, it had been planned to analyze projects under 3 different periods mentioned above. However, the 1999-2003 period was not taken as a separate group because this period is very short compared to the other two and there was not sufficient amount of data to evaluate the effect of exchange rates. As a result, the exchange rates were separated into two periods according to the exchange rate movements and are indicated by the dashed line in Figure 2. In this study, the period of 2003-2010, in which the exchange rates are relatively stable, was defined as the first group, and the fluctuating period of 2010-2018 was defined as the second group. It is considered that the grouping of the projects according to two periods mentioned above will also help to understand the effect of exchange rate and to give more accurate results.

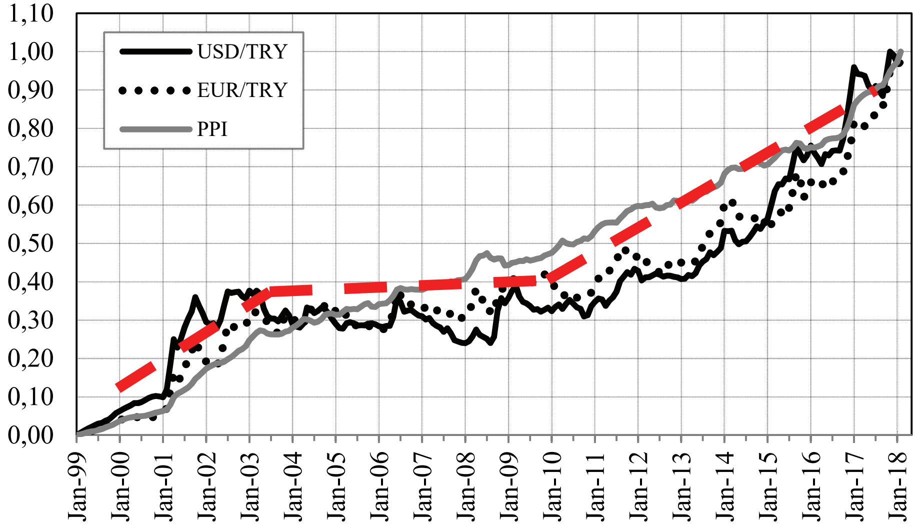


Figure 2 - Normalized USD/TRY, EUR/TRY ER and the PPI between 1999 -2018

In Figure 2, the variation of PPI by years is also demonstrated. When the trend of the index is examined, it can be seen that PPI has a tendency to increase continuously similar to the exchange rates.

The number of non-working days signifies the number of days that cannot be worked due to climatic conditions in construction projects. In order to be able to investigate the effect of the time component, which is one of the two main components of the BTC model, on the completion of the project, the number of non-working days is included in the analysis. The number of non-working days (NWD) is determined by the Ministry of Environment and Urbanization for each province in Turkey [44]. Using this parameter, a grouping was made according to non-working days determined within the scope of the circular published by the Ministry according to conditions of the provinces where the projects take place.

4. ANALYSIS AND RESULTS

4.1. Basic Analysis

Figure 3 shows the time-cost relationship of the building groups according to their intended use. The vertical axis defines the time data in calendar days, while the horizontal axis represents the project costs in Million Turkish Liras. Dashed lines represent +/- one standard deviation values for the data sets. Standard deviation ranges have been found to be higher in hospital and industrial building projects as compared to other groups. This is due to the fact that whilst there are hospital projects costing approximately 1 million TL, there are also hospital projects that cost 164 million TL in the data set. This situation obviously causes the standard deviation to rise especially in terms of the hospital and industrial projects.

The BTC model coefficients for all projects and individual groups are shown in Table 3 and the time-cost curves are shown in Figure 3. There is a strong correlation between the time and cost for all groups with R^2 values ranging from 0.74 to 0.87. However, the results indicate that social housing projects differentiate from other projects with somewhat lower R^2 value (0.54). The project durations of these projects do not increase at the same pace when the curve of social housing projects costs are higher than 30 Million TL, adjusted value. This is due to the fact that social housing projects are composed of multiple buildings and the duration is mostly influenced by the building having the longest construction period within the project.

Table 3 - BTC Model Analysis Results According to Project Types (S=460)

Project Type	K	B	R^2
Social Housing	346.93	0.102	0.54
Individual Building	307.19	0.156	0.79
Hospital	159.57	0.362	0.87
Educational Building	196.43	0.252	0.82
Industrial Building	151.83	0.269	0.74
All Projects	230.04	0.210	0.60

When the projects are analyzed according to their completion times, K coefficients obtained for industrial and hospital projects differ from other projects. However, when B coefficients are examined, it is observed that the hospital projects (B=0.362) have a higher value than industrial projects (B=0.269). As B coefficient is a constant that indicates the sensitivity of time performance to cost level, it has been concluded that while costs of hospital projects, with the highest B value, have increased, the construction periods of these projects have increased more than the other project groups. In Figure 3, it has been seen that with the increase in cost, construction times of hospital projects reach approximately1000 calendar days and differentiate from other projects.

The main reason for the decomposition of the hospital projects on the graph is that the cost of hospital projects increases and the construction time tends to increase more rapidly compared to the other groups. Particularly when the data of the hospital projects are examined in detail, it is seen that the realization period of the projects mostly exceeded the contract period. The curves of the groups are also compared in Figure 4. As can be seen from the figure the hospital projects exhibit a different trend at higher project values compared to the other groups due to their high *B* coefficient. On the other hand, one of the noteworthy findings obtained in this study is that for the other project groups BTC curves demonstrate a very similar trend.

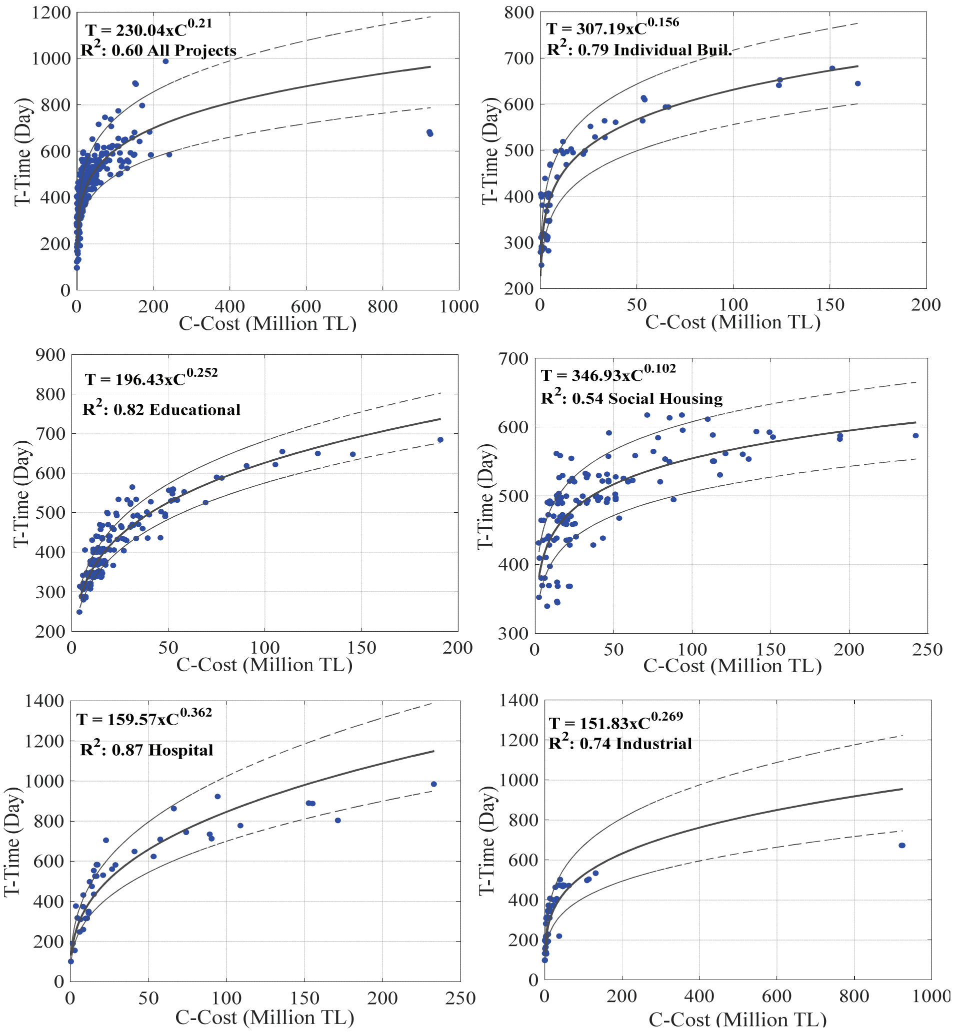


Figure 3 - BTC Model Analysis Results for All Projects

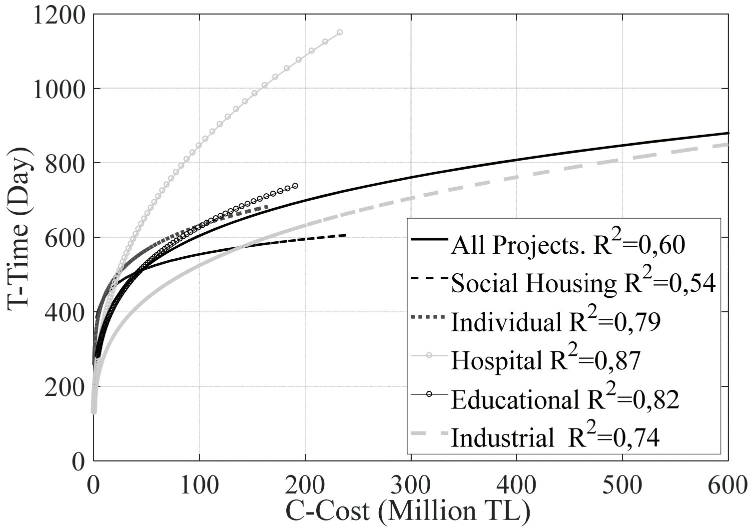


Figure 4 - BTC Model Curves According to Project Groups

As mentioned earlier Bayram (2017) studied time-cost relationships of public buildings in Turkey. While 82% of his data set consisted of educational buildings, 11% of the data set was hospital projects [3]. The equations obtained for the projects of the educational building (Equation 6), the hospitals (Equation 7) and the equation proposed by Bayram (2017) (Equation 5) can be compared as follows;

$$T = 209C^{0.350} \tag{5}$$

$$T = 196.43C^{0.252} \tag{6}$$

$$T = 159.57C^{0.362} \tag{7}$$

When *K* coefficients of this study and Bayram's are compared, Bayram's coefficient (2017) is higher. When the *B* coefficients that shows the slope of the curve are examined, the obtained coefficient in the study of Bayram (2017) is *B* = 0.350 for hospital projects, while the obtained coefficient is *B* = 0.362 in this study for the same project type. As a result, although coefficient *K*, obtained from Bayram (2017), is similar to coefficient *K* for the educational buildings in this study, coefficient *B*, obtained from Bayram (2017)'s study is similar to coefficient B of hospital projects in this study. It is worthwhile to mention that in Bayram's (2017) data set is a composition of 530 public projects which mainly consist of educational buildings (432) and hospital projects (56), therefore direct comparison with this study will not yield accurate results, since all the projects are analyzed together [3].

4.2. BTC Model Analysis According to Different Parameters

Apart from the groupings made in the previous section, the data was further analyzed for non-working days (3 groups) and Exchange Rate (2 groups). The time-cost relationship analysis results of the BTC model for the five building groups, examined with the non-working days and exchange rate parameters considered are summarized in Table 4 and presented in detail in the following sub-sections.

Table 4 - Results of BTC Model Analysis According to NWD and ER Parameters, (S=460)

Project Groups	Parameter	Group 1			Group 2			Group 3		
		K	***B***	***R²***	***K***	***B***	***R²***	***K***	***B***	***R²***
All Projects	Number of the Non-Working Days	261.9	0.165	0.56	177.5	0.269	0.64	254.2	0.197	0.65
Social Housing		343.0	0.103	0.59	403.6	0.059	0.32	330.9	0.116	0.58
Individual Building		312.7	0.147	0.60	305.4	0.160	0.89	277.5	0.179	0.90
Educational Building		190.6	0.255	0.84	170.2	0.334	0.81	213.0	0.247	0.87
Hospital		191.3	0.299	0.83	133.7	0.494	0.98	149.2	0.398	0.85
Industrial Building		195.3	0.198	0.50	144.5	0.275	0.78	249.0	0.158	0.94
All Projects	Exchange Rates	217.0	0.228	0.70	234.8	0.203	0.56			
Social Housing		302.6	0.139	0.67	369.9	0.085	0.50			
Individual Building		-	-	-	307.2	0.156	0.79			
Educational Building		197.3	0.250	0.84	188.7	0.266	0.77			
Hospital		159.1	0.737	0.90	158.8	0.355	0.84			
Industrial Building		223.4	0.172	0.92	135.6	0.313	0.73			

4.2.1. The Effect of Non-Working Days on Time-Cost Relationship of All Projects

In construction projects, especially due to adverse weather conditions such as rain, snow and humidity, there can be non-working (NWD) days within the contract period [44]. Thus, the non-working days reflect the climatic characteristics of the geographical zones. The effect of non-working days on time-cost relationship for the projects in the data set were analyzed according to the circular published by the Ministry of Public Works and Settlement on 07/07/1982 and with the number B-01/İh.İş.Mr.Gr.3/99-2/101156-B [41].

The circular assigned a NWD value for each city as 0, 60, 75, 90, 105, 120, 135, 150 or 165 days. For the purpose of observing regional effect, the number of non-working days has been examined in 3 groups: the cities with 0 and 60 NWD values are in the Group 1, while 75, 90 and 105 NWD values are in Group 2, and the remaining are in Group 3.

The results obtained for each group are shown in Figure 5. One would expect that the projects in Group 3, with the highest NWD, will be realized in the longest time. However, the coefficient *B* (0.269) obtained for the Group 2 projects is higher than the other groups, which means that the project completion times will be longer with increasing costs in Group 2 projects. The projects in the Group 3 are completed in a longer period than the projects in Group 2 up to the limit of TL 150 Million. However, it is understood that the projects in

Group 2 costing 150 Million TL and more will be completed longer than the projects in the Group 3. It is seen in Figure 5 that the projects in Group 1, which have the least non-working days (0-60 days), are completed in a shorter period than the other groups. Since the coefficient *B* is lower than the other two groups, it is seen in the related graph that the increasing cost has less effect on the total number of days for the Group 1.

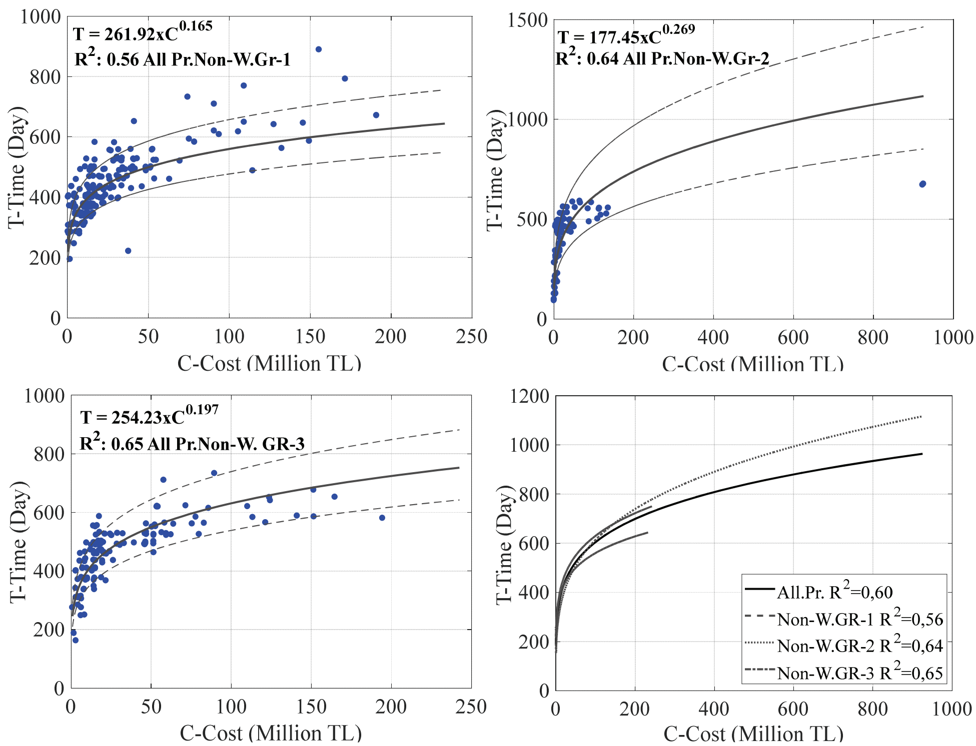


Figure 5 - All Projects- BTC Model Curves According to NWD

4.2.2. The Effect of Non-Working Days on Time-Cost Relationship of Social Housing Projects

Social housing projects were analyzed according to the non-working days and the results obtained are shown graphically in Figure 6.

For the projects in Group 2, the determinant of time-cost relationship (R^2 = 0.32) is lower than the determinant of time-cost relationship (for Group 1 R^2 = 0.59 and for Group 3 R^2 = 0.58) obtained for the other groups is lower. Additionally, Group 2 has a quite large standard deviation than other groups because of the fact that social housing projects in the Group 2 have time and cost values that are highly diverse from each other as seen in Figure 5. Social housing projects carried out in the provinces with the highest number of non-working days, are completed in a longer period than in the provinces with the least number of non-working days (Group 1, 0 and 60 days). However, no significant difference in project completion times was determined for both of the two social housing project groups. The most important reason for this situation can be interpreted, as social housing projects are very similar in structure and completed with a certain systematic method. In terms of all three groups, projects with the cost of approximately 35 million TL can be completed in about 500 days on the average.

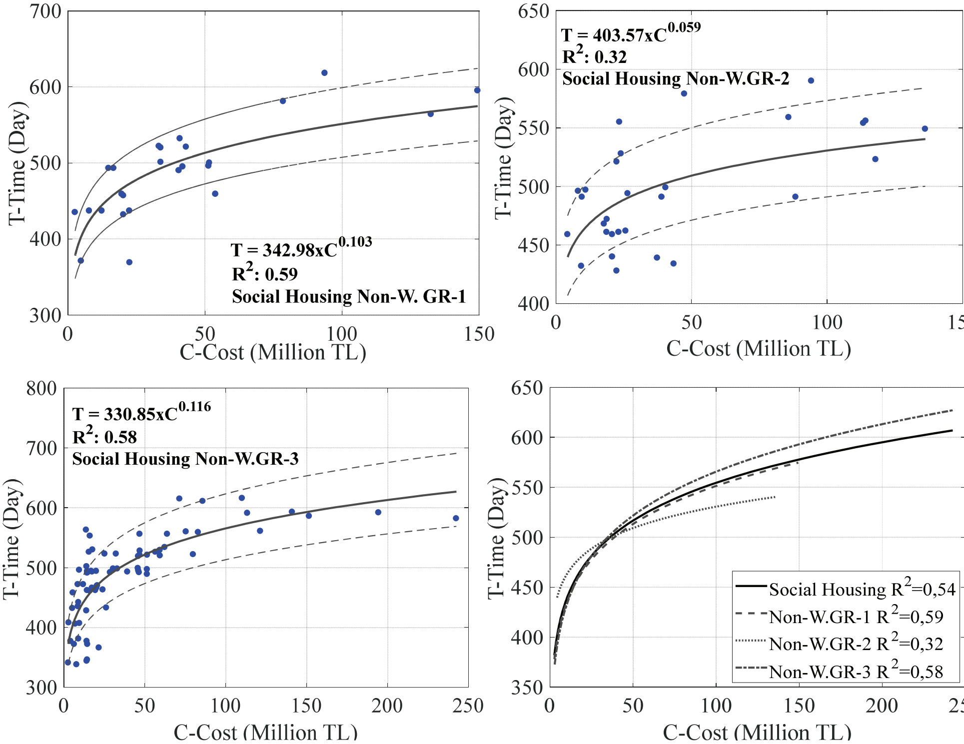


Figure 6 - Social Housing Projects - BTC Model Curves According to NWD

4.2.3. The Effect of Non-Working Days on Time-Cost Relationship of Individual Building Projects

The individual building projects were analyzed according to the non-working days and the results obtained are shown in Figure 7. According to the results, in the first analysis of the individual building projects without the effect of non-working days, determinant of the time-cost relationship of the BTC model was found $R^2 = 0.79$. However, in the analysis with the non-working days parameter, the time-cost relationship for individual building projects in the Groups 2 and 3 (for group 2, $R^2 = 0.89$ and for Group 3, $R^2 = 0.90$) was found to increase. For the Group 1, the time-cost relationship was found to be lower ($R^2 = 0.60$). When the BTC model curves of the individual building projects separated into 3 different groups according to non-working days parameter are examined, it was determined that there is not a significant difference in terms of time-cost relationship. In this context, it can be understood from Figure 7 that the time-cost relationship of the individual building projects carried out by the private sector is not affected by the number of non-working days.

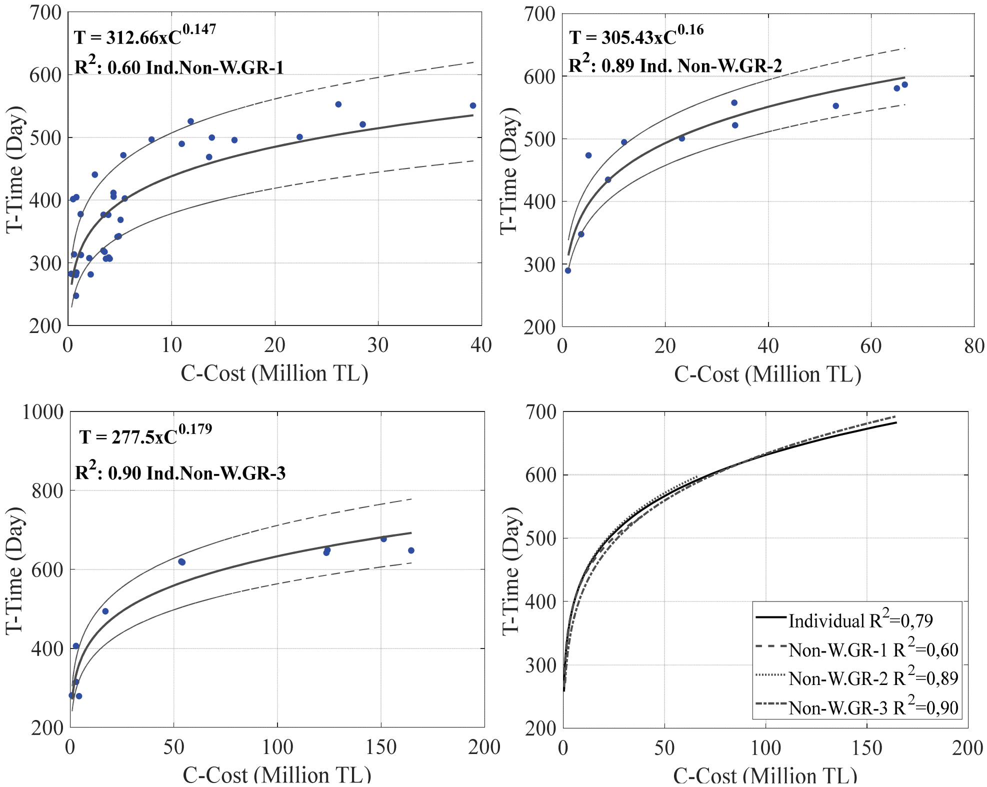


Figure 7 - Individual Building Projects - BTC Model Curves According to NWD

4.2.4. The Effect of Non-Working Days on Time-Cost Relationship of Educational Building Projects

The educational building projects were analyzed according to the non-working days and the results obtained are shown in Figure 8. According to this analysis, the time-cost relationship for all groups was found to be 81% and above. The coefficient *B* (0.334) of the projects in the Group 2 is higher than those of the other groups as seen in the Figure 8.

In this context, it is understood that if the high-cost educational building projects are built in the second group, the projects will be completed longer than those in the Groups 1 and 3. The projects in the Group 1, which have the least non-working days, received close values for the *K* (196,43) and *B* (0,252) coefficients to the coefficients (K = 190,57, b = 0,255) obtained for all educational building projects before grouping. Projects in Group 3 have the highest determination coefficient (R^2 = 0.87) which explains the time-cost relationship being higher than those of other two groups. However, according to analysis after grouping is made according to non-working days, it can be understood that the projects in the Group 1 (0-60 days) can be completed in a shorter period than the other groups.

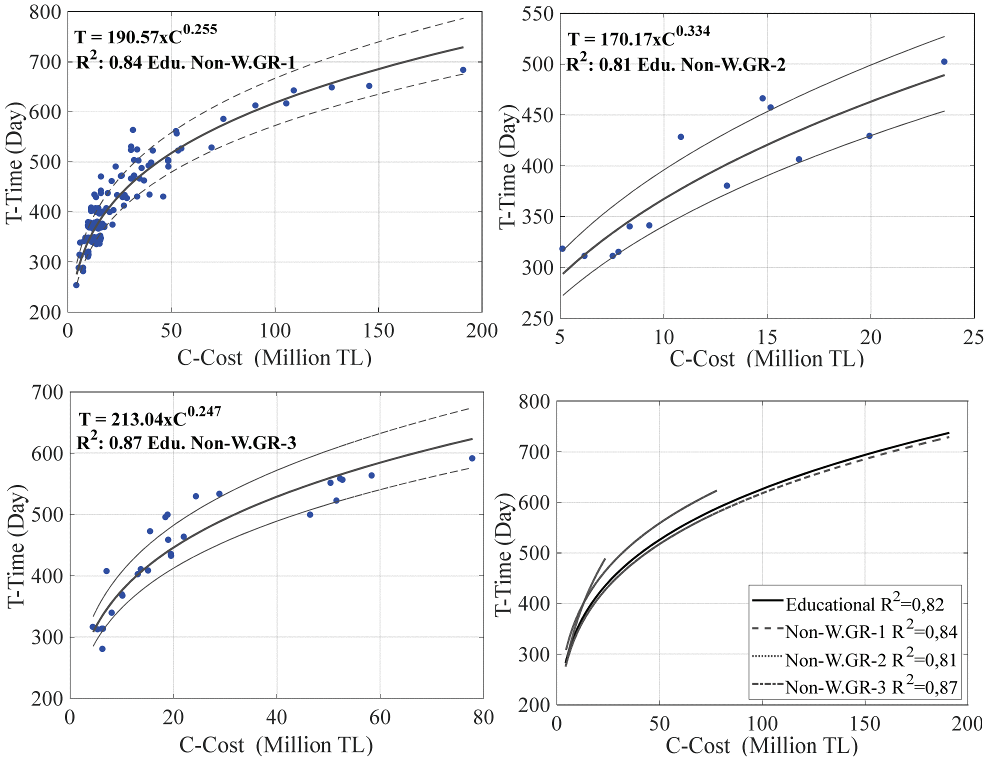


Figure 8 - Educational Building Projects - BTC Model Curves According to NWD

4.2.5. The Effect of Non-Working Days on Time-Cost Relationship of Hospital Projects

The results of the analysis for hospital projects according to the non-working days are shown in Figure 9. According to the results, the determination coefficient levels of BTC models (Group 1 R^2 = 0.83; Group 2 R^2=0.98; Group 3 R^2=0.85) obtained from analysis after grouping, are similar to the coefficient (R^2=0.87) for hospital projects before grouping according to non-working days. It can be understood from Figure 9 that the projects in the Group 1 can be completed in a shorter time than the projects in the Group 3.

In this context, in the case of hospital projects, it is understood that as long as the number of non-working days increases, the average completion times of the projects increases. On the other hand, the time-cost relationship level (98%) of the projects in the second group is obtained significantly higher than the other groups. However, it is important to note that the number of projects in this group was limited to four projects, which in turn may not yield accurate results.

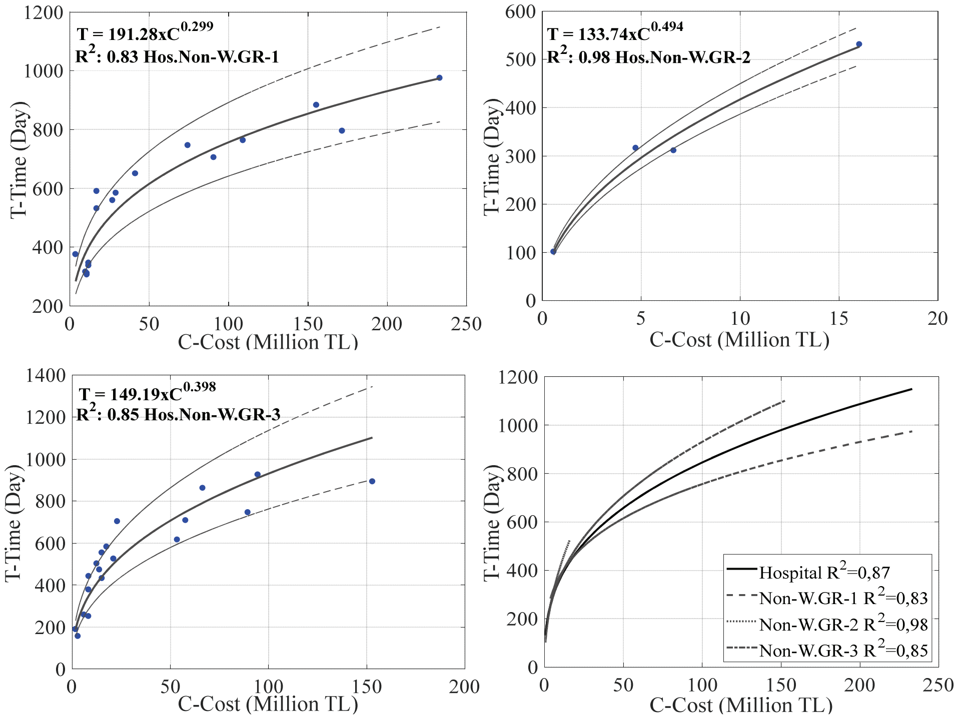


Figure 9. Hospital Projects - BTC Model Curves According to NWD

4.2.6. The Effect of Non-Working Days on Time-Cost Relationship of Industrial Building Projects

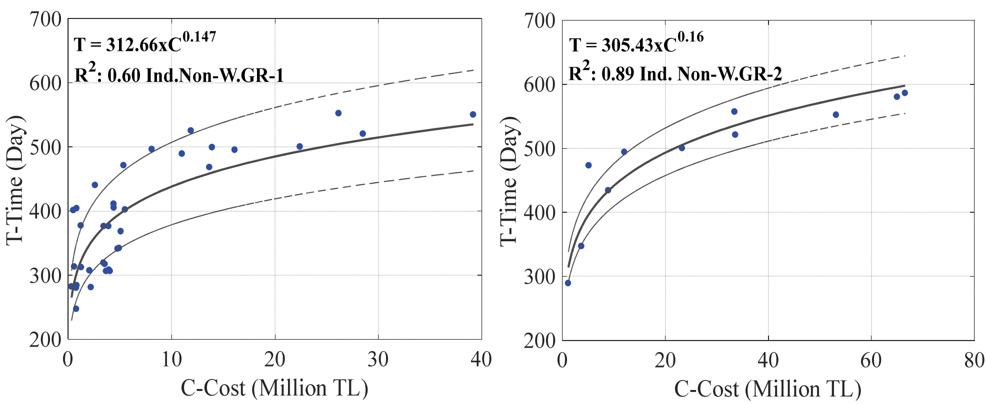


Figure 10 - Industrial Building Projects - BTC Model Curves According to NWD

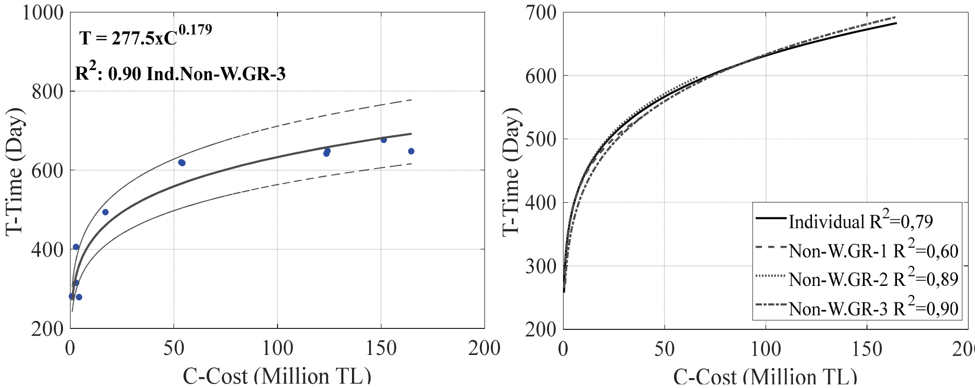


Figure 10 - Industrial Building Projects - BTC Model Curves According to NWD (continue)

The analysis results for the industrial building projects according to the non-working days parameter are shown in Figure 10. Accordingly, strong time-cost relationship was identified in all projects except the projects in the Group 1. The BTC model curves of the industrial projects in Group 2 and 3 seem to be similar. However, the projects in the Group 1 are separated in the graph with their longer construction time. In this context, for industrial building projects, it is seen that as the number of non-working days increases, the completion time of the projects seems to increase.

4.2.7. The Effect of Exchange Rates on Time-Cost Relationship of All Projects

The results of the analysis obtained for all the projects after the grouping according to the exchange rates parameter are shown in Figure 11. According to these results, the time-cost relationship level of projects carried out in the period 2003-2010, when the exchange rates were more stable, is determined as $R^2 = 0.70$, on the other hand, in the period 2010-2018,

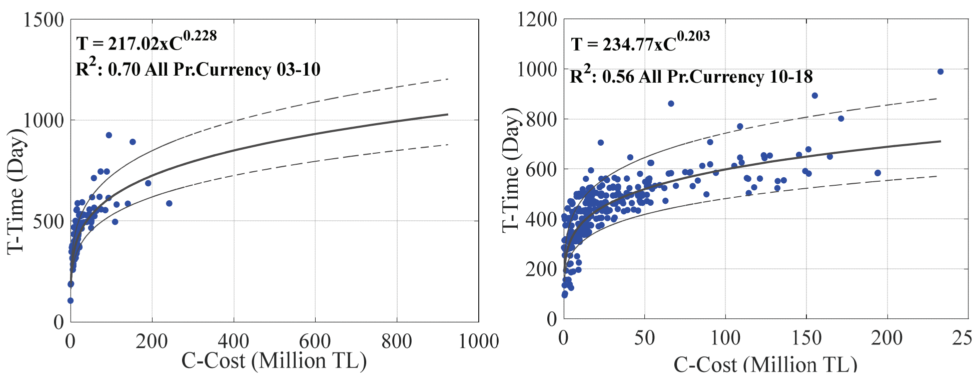


Figure 11 - All Projects - BTC Model Curves According to Exchange Rates

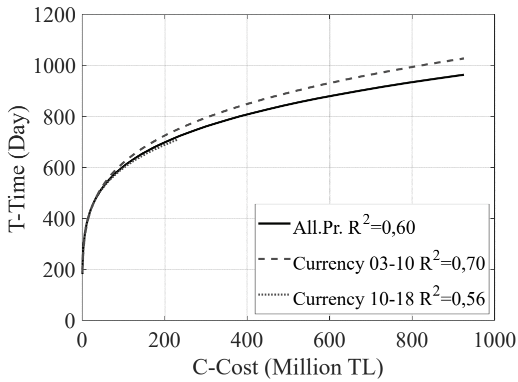


Figure 11 - All Projects - BTC Model Curves According to Exchange Rates (continue)

when the exchange rates were in an increasing trend, R^2 is determined as 0.56. This shows that time-cost relationship of projects in the first period can be explained with more validity than in the second. However, when the effect of the exchange rates parameter is examined in terms of project completion times, it can be seen that the projects are completed in a shorter period in the period of 2010-2018. Especially, in the case of the project costs that are higher than 50 million TL, it is observed that the construction durations of the projects during the period between years 2010-2018 are less than those in the projects carried out in the previous period.

4.2.8. The Effect of Exchange Rates on Time-Cost Relationship of Social Housing Projects

The results obtained for social housing projects after the grouping according to the exchange rates parameter are shown in Figure 12. The level of the time cost relationship for the projects carried out in the period 2003-2010 is $R^2 = 0.67$, while it is $R^2 = 0.50$ for the period 2010-2018. Projects in Group 1, up to the limit of 40 Million TL cost, can be completed in a shorter time than the other groups, while projects in the Group 2 can be completed in a shorter time when the costs exceed 40 Million TL.

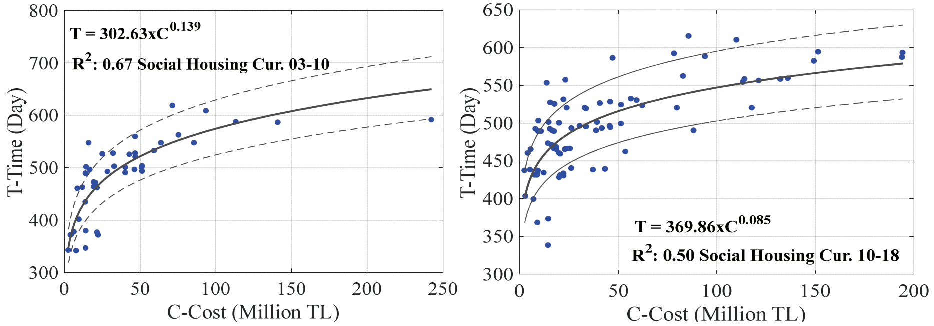


Figure 12 - Social Housing Projects - BTC Model Curves According to Exchange Rates

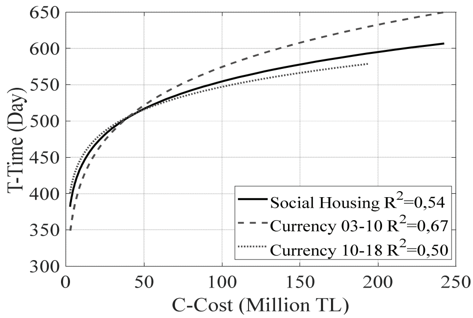


Figure 12 - Social Housing Projects - BTC Model Curves According to Exchange Rates (continue)

4.2.9. The Effect of Exchange Rates on Time-Cost Relationship of Educational Building Projects

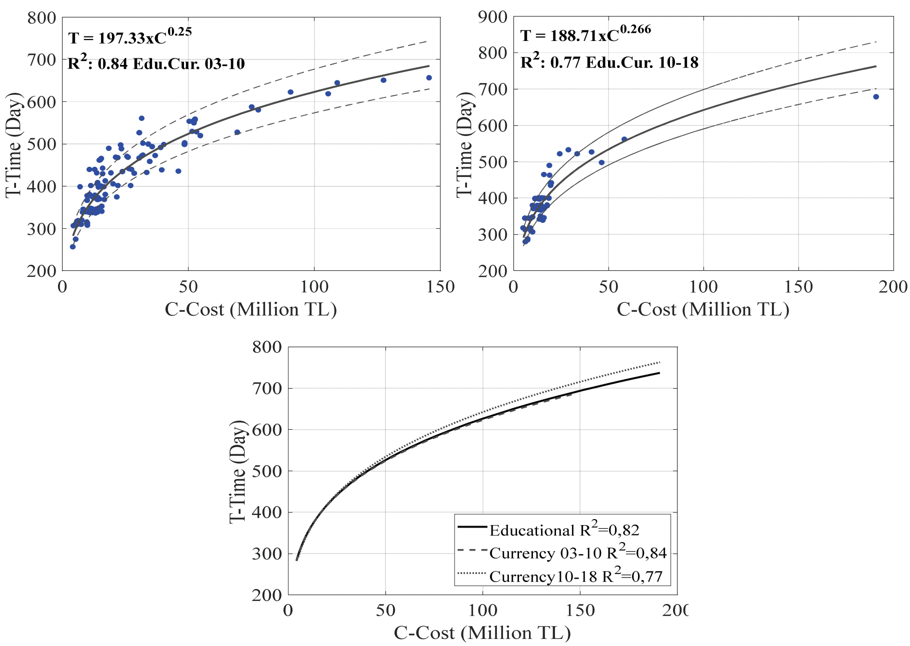


Figure 13 - Educational Building Projects - BTC Model Curves According to Ex. Rates

The time-cost relationship of educational building projects under the exchange rate effect has been examined and the analysis results are shown in Figure 13. Accordingly, it seems that the BTC model explains the time-cost relationship in both two periods. As seen in the graphics, the time-cost relationship level of the educational building projects realized in the period of 2003-2010 are at a higher level than that of the projects of 2010-2018 period. It can also be concluded that the educational building projects carried out in the period of 2003-2010 were completed in a shorter period than the educational building projects built between 2010 and 2018.

4.2.10. The Effect of Exchange Rates on Time-Cost Relationship of Hospital Projects

The time cost relationship of hospital projects under the exchange rate effect is shown in Figure 14 as a result of the analysis. The projects in both groups, which are formed according to the exchange rate parameters, did not show any distinction until the cost of approximately of 25 Million TL. However, it can be understood from Figure 14 that the projects, with the cost of over 25 Million TL, can be completed in a shorter period from 2010 to 2018 than in the first period. As a result of this analysis, the higher R^2 values obtained for both periods show that the time-cost relationship level is higher.

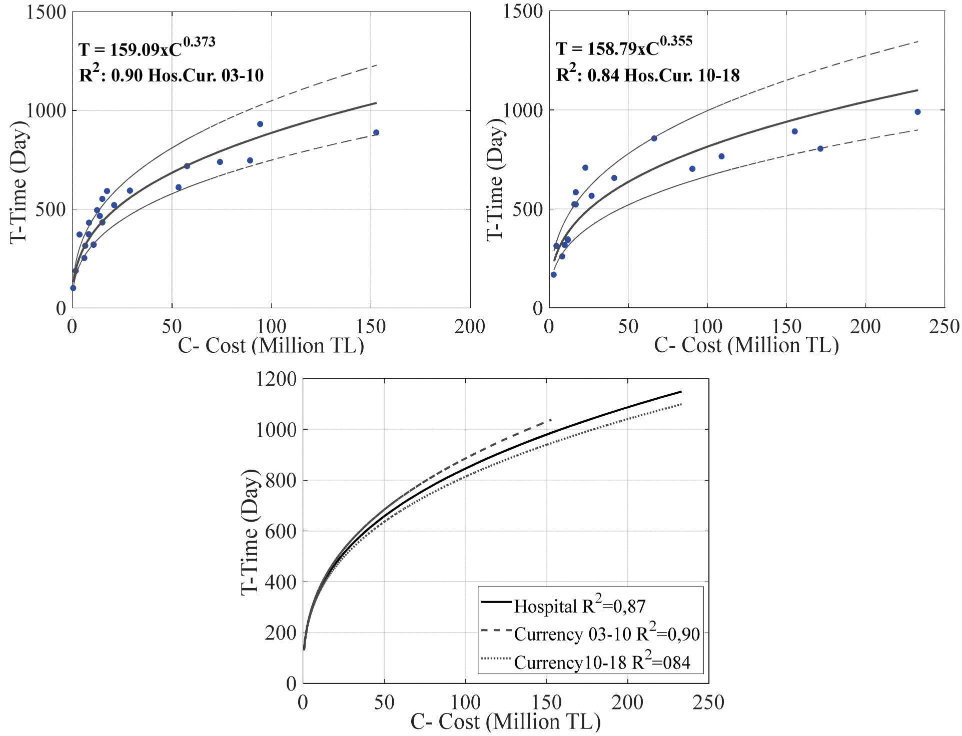


Figure 14 - Hospital Projects - BTC Model Curves According to Exchange Rates

4.2.11. The Effect of Exchange Rates on Time-Cost Relationship of Industrial Building Projects

The time-cost relationship of the industrial building projects, under the exchange rate effect, has been examined and the analysis results are shown in Figure 15. Industrial building projects in Group 1 have a more significant time-cost relationship like in the other type of projects when compared to projects built in 2010-2018 period. However, unlike other building groups, industrial projects are completed in a shorter period in the period when foreign exchange is stable (2003-2010). On the other hand, in the period of 2010-2018 when the foreign exchange is in a rising trend, the industrial projects can be completed considerably longer. Companies that are operating in the industrial sector in our country are also vulnerable to exchange rates, and this situation may lead to a deterioration of the payment balance of infrastructure investments [45]. On the other hand, the construction cost of the industrial building projects are basically a function of the main construction materials used, namely reinforcing steel, structural steel and concrete [46]. Whereas, the prices of these materials are all indexed to the exchange rates.

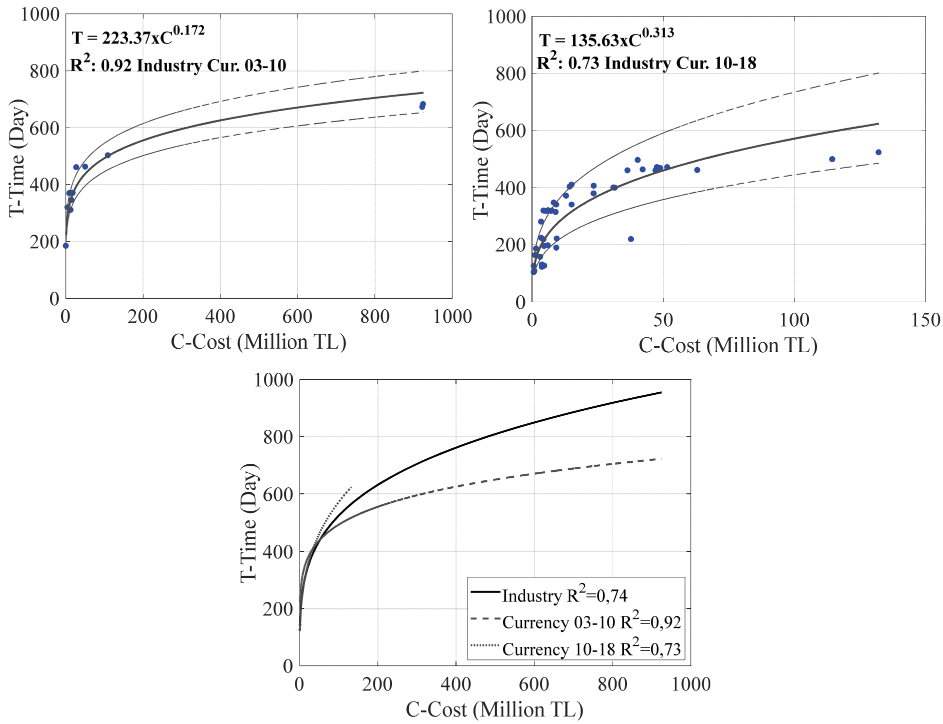


Figure 14 - Industrial Building Projects - BTC Model Curves According to Exchange Rates

4.2.12. The Effect of Exchange Rates on Time-Cost Relationship of Individual Building Projects

The grouping with the foreign exchange parameter could not be applied for individual building projects group since the data of the building projects are distributed only between the years 2010-2018. In this context, there could be no possibility of comparison between two periods in order to reveal the effect of exchange rates on the time-cost relationship of these projects.

5. DISCUSSION

The effects of the number of non-working days and exchange rate parameters on time-cost relationship of the Project groups are summarized in Table 4. As a result of the analysis carried out according to non-working days, the time-cost relationship levels in all groups for all other projects except social housing projects are significantly explained in the 59-98% range. On the other hand, in the period of 2010-2018, it is seen that the level of explanation of time-cost relationship of the BTC Model is lower than in the period of 2003-2010 due to the increase of the exchange rates against Turkish Lira. In this context, it has been determined that the time-cost relationship is also influenced by the fluctuation of the exchange rate, since this fluctuation affects project costs.

The fact that the time-cost relation between the periods of 2003-2010, when the foreign exchange is stable, is higher compared to the period 2010-2018 when the foreign exchange is fluctuating shows the effect of the exchange rates on the Turkish Construction Sector. The reason for this is the fact that the equipment used in the construction sector and some of the raw materials are imported with using foreign currencies [41]. Similarly, it is understood that the number of non-working days is an important factor affecting the time-cost relationship.

The higher the tender prices affect hospital projects at the highest level in terms of duration. As the project cost increase for these building groups, the delivery time is also increasing remarkably. On the other hand, social housing projects with the cost of 50 million TL and above, are seen to have a shorter period of completion than the other building groups. Despite the high tender prices in this type of construction projects, delivery times are not affected at the same time. When compared to other building groups, the BTC model has the lowest level of determination (R^2 = 0.56) for the time-cost relationship for social housing projects. However, it has been found that the level of time-cost relationship in the projects built by the private sector is higher than that of the projects built by TOKİ (Public). It is found that the time-cost relationship for this study varies for both the public and private sectors, according to the results the performance of private projects is better than the public in this study as Le-Hoai et al. (2009) [34] revealed the same conclusion in their respective study. Some of the reasons for the low performance of public projects may change orders related to the projects, such as concept changes, design problems, and other unidentified causes as Albalushi et al. (2013) mentioned in his study [47]. According to Kaka and Price (1991), the reason for the lower performance of public projects than the private projects is the high level of the interest and concern of the private sector customers on the time and the cost of the projects [12]. This is a very critical and sensitive issue for investors to have the projects start the operation phase within the planned budget as soon as possible. The results of this study also support the results of Kaka and Price (1991).

The effects of the non-working days on the time-cost relationship is analyzed in this study and it has been revealed that the completion times of the projects are shorter in the cases where the non-working days are the lowest (Group 1, 0-60 days). On the other hand, no remarkable effect of non-working days was observed on the time-cost relationship of the private sector building projects. The results of the analysis have revealed that the industrial projects were directly affected by the exchange rate fluctuation (Figure 15). Projects built between 2010 and 2018, except the industrial building projects, were completed in a shorter period than projects between 2003 and 2010. On the other hand, it is understood from this study that BTC model reveals the time-cost relationship for all project groups at a very high level between 2003 and 2010. The fact that BTC model explains time-cost relationship in the period with stable exchange rates better than the in the period with increasing exchange rates, indicates that the sector is affected by the exchange rate fluctuations

6. CONCLUSION

In this study, the time-cost relationship of the building projects realized in Turkey was investigated based on the Bromilow's time-cost prediction model. The data set used comprised of 460 projects intended for different uses at climatic regions of the country. The time-cost relationship curves of each group, as well as all the projects as a group, were determined taking into consideration the changes in the exchange rate and the number of non-working days (namely the climatic effects). Subsequently, it is concluded that the consideration of the number of non-working days and the exchange rates can have strong impact on the time –cost relationship. Also, the curves can be refined further by using larger data sets.

References

[1] Barrie, D. S., & Paulson, B. C., Professional construction management: including CM, design-construct, and general contracting. McGraw-Hill Science/Engineering/Math. 1992.

[2] Bromilow, F. J., Contract Time Performance Expectations and the Reality. Building Forum, 1, 70-80, 1969.

[3] Bayram, S.. Duration prediction models for construction projects: In terms of cost or physical characteristics?. KSCE Journal of Civil Engineering, 2017, 21.6: 2049-2060.

[4] Petruseva, S., Pusic, D. C., & Pancovska, V. Z. Model for Predicting Construction Time by Using General Regression Neural Network. International Scientific Conference People, Buildings and Environment 2016 (PBE2016) 29 September – 1 October, 2016, Luhacovice, Czech Republic, 2016.

[5] Bayram, S. & Al-Jibouri, S. Reference Class Forecasting Method in Predicting Construction Project Duration. Süleyman Demirel University Journal of Natural and Applied Sciences, 22 (3), pp.1132-1142, 2018

[6] Ariful Haque, K.M., & Hasin, M.A.A. Genetic algorithm for project time-cost optimization in fuzzy environment. Journal of Industrial Engineering and Management, 5(2), pp.364-381, 2012.

[7] Odabaşı, E., Models for Estimating Construction Duration: An Application for Selected Buildings on the METU Campus. Ankara: Ortadoğu Teknik Üniversitesi Uygulamalı Bilimler Enstitüsü, 2009.

[8] Love, P. E. D., Tse, R. Y. C., & Edwards, D. J., Time-cost relationships in Australian building construction projects. Journal of Construction Engineering and Management (ASCE), Vol. 131, No. 2, pp. 187-194, 2005.

[9] Ireland, V., The Role of Managerial Actions in the Cost Time and Quality Performance of High-Rise Commercial Building Projects. Construction Management and Economics, 3(1), 59-87, 1985.

[10] Choudhury, I. & Rajan, S. S. Time-Cost Relationship for Residential Construction in Texas. CIB, 2003.

[11] Ng, S., Mak, M., Skitmore, R. & Varnam, M. The Predictive Ability of Bromilow's Time-Cost Model. Construction Management and Economics, 19(2), 165-173, 2001.

[12] Kaka, A., & Price, A. D., Relationship between value and duration of construction projects. Construction Management and Economics, 9(4), 383-400, 1991.

[13] Kumaraswamy, M., Chan, D., Determinants of Construction Duration Construction Management and Economics, 13(3), 209-217, 1995.

[14] Hoffman, G. J., Thal Jr., A. E., Webb, T. S. & Weir, J. D., Estimating Performance Time for Construction Projects. 23, 193-199, 2006.

[15] Chan, A., Time-Cost Relationship of public sector projects in Malaysia. International Journal of Project Management, 19, 223-229, 1999.

[16] Endut I.R., Akintoye, A. & Kelly, J., Relationship Between Duration and Cost of Malaysian Construction projects", Proceedings of the International Conference in the Built Environment in the 21st Century, Kuala Lumpur (ICiBE). 2006.

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

[18] Ojo, G., A Study of the Relationship Between Contract Period and Tender Sum of Construction Projects in South-Western Nigeria. M.Sc. Thesis, Obafemi Awolowo University, Ile-Ife. 2001.

[19] Bromilow, F.J. & Henderson, J.A., Procedures for reckoning and valuing the performance of building contracts. The Chartered Builder, 10(9), 57. 1976.

[20] Yeong, C., Time and Cost Performance of Building Contracts in Australia And Malaysia. University of South Australia. M.Sc. Thesis, University of South Australia School of Building and Planning, 1994.

[21] Mizanur, R. M., Dai, L. Y., Khanh, H. D. & Hyun, C. Y., Development of Time-Cost Models for Building Construction Projects in Bangladesh. 3 (4, Issue 3,). Korea: Korean Institute of Construction Engineering and Management, 2014.

[22] Zujo, V. & Pusic, D. C., Application of "Time-cost" Model in Construction Project Management. Abstracts of the 8th International Conference Organization, Technology and Management in Construction and 5th SENET Conference (s. 1-7). Umag, Hrvatska, 2008.

[23] Petruseva, S., Zujo, V. & Pancovska, V. Z. Neural Network Prediction Model for Construction Project Duration. International Journal of Engineering Research & Technology (IJERT), 2(11), 1646-1654. 2013.

[24] Sun, C., Xu, J., Estimation of Time for Wenchuan Earthquake Reconstruction in China. Journal of Construction Engineering and Management, 137(3), 179-187, 2010.

[25] Sriana, T. & Hayati, K., Time-Cost Relationship Model on the Construction of Education Building in Aceh Province. Journal of Asian Scientific Research, 5(7), 328-339, 2015.

[26] Ameyaw, C., Mensah, S. & Diss, Y. Applicability of Bromilow's Time-Cost Model on Building Projects in Ghana. 4th West Africa Built Environment Research (Waber) Conference, Abuja, Nigeria, 24-26, 2012.

[27] Le, Long, L.-H. Y., Bromilow's Time-Cost Equation for Building Projects in Korea. Korean Society of Civil Engineers, (7) 812-815, 2009.

[28] Choudhury, I. Time-Cost Relationship for Infrastructure Projects in India. (s.14.1270.1). American Society for Engineering Education, 2009.

[29] Car-Pušić, D. & Radujković, M., Model for Proposed Construction Time for Transportation Facilities. Suvremeni Promet, 30(1/2), 127-130, 2010.

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

[31] Waziri, B. & Nurudeen, Y. Application of Bromilow's Time-Cost (BTC) Model to Highway Projects in Nigeria. Arid Zone Journal of Engineering, Technology and Environment, 10, 1-11, 2014.

[32] Mačková, D. & Bašková R., Applicability of Bromilow´s Time-cost Model for Residential Projects in Slovakia, SSP - Journal of Civil Engineering Vol. 9, Issue 2,5-12, 2014.

[33] Özçekiç, A. F., Türk Yüklenicilerin Uluslararası Teklif Aşamasında Kullanabileceği Bir Süre Maliyet Modeli, Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi Fen Bilimleri Enst. Bilim ve Teknoloji Enstitüsü., İstanbul, 2007.

[34] Le-Hoai, L., & Dai Lee, Y. Time-cost relationships of building construction project in Korea. Facilities, 27 (13/14), 549-559, 2009.

[35] Okere, G. An Evaluation of a Predictive Conceptual Method for Contract Time Determination on Highway Projects Based on Project Types. International Journal of Civil Engineering, pp.1-17, 2018.

[36] Adeyemi, A. Y., & Masalila, K. Delay factors and time-cost performance of construction projects in Gaborone City Council, Botswana. Journal for the Advancement of Performance Information & Value, 8(1), 1-18, 2016.

[37] ISO 4217, obtained from https://www.currency-iso.org/en/home.html. web address on 03.03.2018.

[38] Martin, J., Burrows, T. K. & Pegg, I. Predicting Construction Duration of Building Projects. XXIII FIG Congress. Munich, Germany, 2006.

[39] Mathworks Matlab R2017a, https://www.mathworks.com/

[40] Kenley, R., Financing Construction: Cash flows and cash farming. Routledge. 2003.

[41] Kılıç, E., & Yıldırım, K., Sektörel reel döviz kuru volatilitesinin Türk imalat sanayi ihracatı üzerine etkileri. Anadolu Üniversitesi Sosyal Bilimler Dergisi, 15(1), pp.13-25, 2015.

[42] TCMB (Central Bank of the Republic of Turkey) Foreign currency information, according to periods, is obtained from https://evds2.tcmb.gov.tr/index.php?/evds/seriemarket web address on 03.03.2018.

[43] Ministry of Environment and Urbanization. Circular issued on 07/07/1982 and numbered B-01 / İh.ş.Mr.Gr.3 / 99-2 / 101156-B, 1982

[44] Thorpe, D., & Karan, E. P., Method for calculating schedule delay considering weather conditions. In Proceedings of the 24th annual conference of the Association of Researchers in Construction Management (ARCOM 2008) (Vol. 2, pp. 809-818). Association of Researchers in Construction Management (ARCOM). 2008.

[45] TOBB, (2011). Türkiye İnşaat Malzemeleri Sektör Görünüm Raporu, 2011.

[46] Dikmen, S. Ü., & Özek, S. Deprem bölgelerinde zemin sınıfının sanayi yapılarının maliyetine etkisi. İMO Teknik Dergi, 22 (108), 5543-5558, pp.2011.

[47] Albalushi, I., Usman, F., & Alnuaimi, A. S., Construction Cost Overrun and Variations: Investigation on its Causes and Consequences. Australian Journal of Basic and Applied Sciences, 7(14), 311-323, 2013.

Teknik Dergi, 2020 9897-9917, Paper 572

Investigating Acceptable Level of Travel Demand before Capacity Enhancement for Signalized Urban Road Networks

Özgür BAŞKAN[1]
Hüseyin CEYLAN[2]
Cenk OZAN[3]

ABSTRACT

Increasing travel demand in urban areas triggers traffic congestion and increases delay in road networks. In this context, local authorities that are responsible for traffic operations seek to strike a balance between traffic volume and capacity to reduce total travel time on road networks. Since signalized intersections are the most critical components of road networks in terms of safety and operational issues, adjusting intersection signal timings becomes an effective method for authorities. When this tool remains incapable of overcoming traffic congestions, authorities take expensive measures such as increasing link capacities, lane additions or applying grade-separated junctions. However, it may be more useful to handle road networks as a whole by investigating the effects of optimizing signal timings of all intersections in the network. Therefore, it would be useful to investigate the right time for capacity enhancement on urban road networks to avoid premature investments considering limited resources of local authorities. In this study, effects of increasing travel demand on Total Travel Cost (TTC) is investigated by developing a bi-level programming model, called TRAvel COst Minimizer (TRACOM), in which the upper level minimizes the TTC subject to the stochastic user equilibrium link flows determined at the lower level. The TRACOM is applied to Allsop and Charlesworths' network for different common origin-destination demand multipliers. Results revealed that TTC values showed an approximate linear increase while the travel demand is increased up to 16%. After this value, TTC showed a sudden spike although the travel demand was linearly increased that means optimizing signal timings must be supported by applying capacity enhancement countermeasures.

Note:

- This paper has been received on September 26, 2018 and accepted for publication by the Editorial Board on March 18, 2019.
- Discussions on this paper will be accepted by May 31, 2020.

- https://dx.doi.org/10.18400/tekderg.464260

1 Pamukkale University, Department of Civil Engineering, Denizli, Turkey - obaskan@pau.edu.tr - https://orcid.org/0000-0001-5016-8328

2 Pamukkale University, Department of Civil Engineering, Denizli, Turkey - hceylan@pau.edu.tr - https://orcid.org/0000-0002-8840-4936

3 Adnan Menderes University, Department of Civil Engineering, Aydın, Turkey - cenk.ozan@adu.edu.tr - https://orcid.org/0000-0003-0690-6033

Keywords: Total travel cost, urban road networks, signal timing optimization, investment decisions.

1. INTRODUCTION

Travel demand has substantially increased during recent years as a consequence of fast growing population and correspondingly generated mobility need especially in developing countries. Since a developing country can be defined as a country with less developed industrial base and less gross domestic product per capita relative to developed countries, urban road network management is a lot more essential for these countries. As developing countries have limited financial resources, they are required to efficiently use their resources especially in urban road network management in which investments made to improve road networks performance are quite expensive. A general view in urban road network management is that the local authorities generally try to take some countermeasures such as increasing link capacities, improving junctions, applying grade-separated junctions etc. on a budget basis to cope with increasing travel demand over time. However, many authorities apply these measures after they came across serious problems in terms of traffic congestion and related issues. Since authorities cannot properly predict the influence of increasing travel demand on the performance (e.g. total travel cost) of road networks, investments to overcome such problems cannot be made timely. Moreover, from the point of steering investment decisions, some investments made to increase network's performance may be unnecessary because authorities are already able to overcome such problems by applying some cost-efficient measures (e.g. setting signal timings) even in case of increasing travel demand. At this point, local authorities need to have information in order to decide whether they make expensive infrastructure investments or use suitable signal timings to overcome problems revealed with increasing travel demand. In other words, authorities can make their investment decisions more accurately if they know how much longer transport supply will respond to the increase in travel demand by applying suitable signal timings. As known, signalized intersections are designed in order to reduce overall cost in a road network and to ensure intersection safety as well. However, using inappropriate signal timings leads to increasing total travel cost in the network and to reducing its capacity. Thus, the problem of finding appropriate signal timings is an important issue for local authorities who are responsible for traffic management. Many researchers have investigated this problem with regards different perspectives as given in Table 1 starting almost as long back as the last half century.

As can be seen in Table 1, studies on traffic signal timing optimization have started with well-known work presented by Webster [1]. Until the mid-1990s signal timing optimization had been carried out by using mathematical methods generally aiming to minimize delay. However, from the beginning of the 2000's, heuristic algorithms started to be used in signal optimization due to their successful applications especially in optimization field and to eliminate negative effects of derivatives of complex mathematical expressions used for optimizing signal timings. Since that time, various objective functions have been used instead of minimizing delay. In recent years, one of the most important issues is the selection of objective function to optimize signal timings in a road network, more properly. In this study, total travel cost (TTC) as an objective function has been selected in order to optimize signal timings. It is obvious that link cost function should be firstly defined to find the TTC in a given road network. For this purpose, the link cost function is presented as the sum of free-flow travel time, average uniform delay, and average random plus oversaturation delay per vehicle in order to define a network travel cost closer to the reality. Another reason for the use of the TTC is that the network-level solution of traffic congestion problem may be more effective in comparison with junction-level solution in which authorities try to solve the problem only for considered single junctions. This approach may increase the network congestion level instead of reducing.

Table 1 - Studies on traffic signal timing optimization

Study	Method	Objective function
Webster [1]	Mathematical methods	Minimizing delay
Allsop [2]	Convex programming	Minimizing delay
Allsop [3]	Linear programming	Maximizing capacity
Wong [4]	Approximate expressions	Minimizing delay
Heydecker [5]	Decomposition approach	Minimizing performance index
Wong [6]	Non-linear programming	Minimizing performance index
Wong [7]	Parallel computing	Minimizing performance index
Wong et al. [8]	Heuristic algorithm	Minimizing performance index
Girianna and Benekohal [9]	Heuristic algorithm	Maximizing number of vehicles
Ceylan and Bell [10]	Heuristic algorithm	Minimizing performance index
Ceylan [11]	Heuristic algorithm	Minimizing performance index
Chen and Xu [12]	Particle swarm optimization	Minimizing performance index
Dan and Xiaohong [13]	Genetic algorithm	Minimizing delay
Li [14]	Cell transmission model	Maximizing number of vehicles
Liu and Chang [15]	Genetic algorithm	Minimizing travel time
Ceylan and Ceylan [16]	Hybrid heuristic algorithm	Minimizing performance index
Dell'Orco et al. [17]	Harmony Search algorithm	Minimizing performance index
Dell'Orco et al. [18]	Bee colony algorithm	Minimizing performance index
Ozan et al. [19]	Reinforcement learning	Minimizing performance index
Christofa et al. [20]	Mixed integer programming	Minimizing person based delay
Srivastava and Sahana [21]	Evolutionary algorithm	Minimizing total wait time
Abdul Aziz et al. [22]	Reinforcement learning	Minimizing average delay

The main objective of this paper is to investigate the acceptable level of travel demand before capacity enhancement for signalized road networks. In this study, increasing travel demand, which is a result of growing population, increasing mobility requirement and changing land use pattern, is represented by a common Origin-Destination (O-D) demand multiplier. The base O-D demand matrix of a road network can be increased by this multiplier in order to find the acceptable level of travel demand, which can be stated as a critical border for capacity enhancement on a road network. In this context, TTC as an objective function is defined as the sum of the multiplication of each link equilibrium flow and corresponding travel cost on a given road network. To minimize the TTC, a bi-level programming model, called **TRA**vel **CO**st **M**inimizer (TRACOM), is presented in which the upper level minimizes the TTC by using equilibrium link flows determined at the lower level. As known, equilibrium link flows can be found depending on either deterministic or stochastic approaches. In this study, Stochastic User Equilibrium (SUE) link flows are determined by using Path Flow Estimator (PFE) developed by Bell et al. [23] in order to represent the users' behaviors against signal timing changes performed at the upper level. The TRACOM model is developed based on Differential Evolution (DE) algorithm framework that is a widely used meta-heuristic method for solving complex optimization problems. Finally, the TRACOM is applied to Allsop and Charlesworths' network for different common O-D demand multipliers in order to reveal the effect of increasing travel demand on TTC.

The paper is organized as follows. Section 2 introduces problem formulation. A bi-level model is given in the third section. Section 4 presents a numerical application. Finally, last section is about discussions and conclusions.

2. PROBLEM FORMULATION

Local authorities are responsible to manage networks in urban roads, and they try to minimize the TTC even in case of increasing travel demand conditions to create robust road networks. On the other hand, minimizing TTC provides benefits for road users by reducing their travel time because they aim to complete their travels within minimum travel time. As known, travel demand continues to increase over the last years due to growing population, and thus increasing travel demand leads to an increase in the TTC in urban road networks. At this point, the problem is that how long transport supply will respond to increasing travel demand by optimizing signal timings. By this way, local authorities will not have to make expensive investments, and they can manage robust road networks with cost-efficient methods.

For this purpose, considering a road network with a set of O-D pairs, K, a set of directed links, A, a set of paths, R, and a set of nodes, N, link cost function on link a can be expressed as given in Eq. (1).

$$c_a(q_a) = c_a^0 + d_a^U + d_a^{ro}(t) \tag{1}$$

where q_a is flow on link a, $a \in A$, c_a^0 is free flow travel time on link a, d_a^U is average uniform delay to a vehicle arriving on link a (i.e. uniform component of total delay), $d_a^{ro}(t)$ is average random plus over saturation delay to a vehicle arriving on link a at time slice t (i.e. random component of total delay).

2.1. Random Plus Oversaturation Delay Component

Assuming that time slice t equals time period T, $d_a^{ro}(T)$ can be expressed by using well-known traffic simulation software TRANSYT delay formula proposed by Vincent et al. [24] as follows:

$$D_a^{ro}(T) = \frac{T}{4}[((q_a - \mu_a)^2 + \frac{4q_a}{T})^{0.5} + (q_a - \mu_a)] \quad (2)$$

$$d_a^{ro}(T) = \frac{D_a^{ro}(T)}{q_a} \quad (3)$$

where $D_a^{ro}(T)$ is random plus oversaturation delay on link a for time period T, $d_a^{ro}(T)$ average random plus oversaturation delay to a vehicle arriving on link a for time period T, and μ_a is the capacity for link a.

2.2. Uniform Delay Component

The calculation for uniform component of total delay for each link can be carried out on the basis of whole cycles for uniform arrivals and departures. On the other hand, delay and number of queues which are calculated over the time slice t is based on whole cycles which begin and end at the starts of effective red times. Thus, the uniform component of delay for each link a can be defined in two ways according to the degree of saturation of each link in the network as follows:

(*i*) For oversaturated links with $x_a \geq 1$;

(*ii*) For undersaturated links with accumulated queues $x_a < 1$ which can be identified as those with $L_a^{ro}(t) > L_a^s$.

where $L_a^{ro}(t)$ is random plus oversaturation component for the number of queueing vehicles at time slice t, L_a^s is number of queueing vehicles in steady state, and x_a is the degree of saturation on link a.

For oversaturated links: Let L_a^U be uniform queue, D_a^U be uniform delay, d_a^U be delay to a vehicle, Λ_a be proportion of green to cycle time, and C be the cycle time, the effect of cyclic variation in the cumulative arrivals, q_a on L_a^U and D_a^U is assumed as zero [25]. Thus, L_a^U and D_a^U are calculated on the basis of the difference between the cyclic cumulative departure graph from TRANSYT and uniform departure rate (i.e. capacity), μ_a for each link a in the time period T according to following expressions (see for details [26]).

$$L_a^U = \frac{C\mu_a(1-\Lambda_a)}{2} \tag{4}$$

$$D_a^U = L_a^U \tag{5}$$

$$d_a^U = \frac{C(1-\Lambda_a)}{2} \tag{6}$$

For undersaturated links with accumulated queues: Considering that there may be queue in undersaturated links, if the initial queue length, $L_a^{ro}(0)$ (i.e. t=0) is higher than the steady state queue length, L_a^s for time slice t, time is required for the accumulated queue of link a at the start of t to be dissipated. Let τ_a be the time needed, the following expression can be used [27].

$$\tau_a = \frac{L_a^{ro}(0) - L_a^s}{\mu_a(\hat{x}_a - x_a)} \tag{7}$$

where $\hat{x}_a$ is the degree of saturation for which the equilibrium queue is equal to the initial queue. $L_a^{ro}(0)$ and $\hat{x}_a$ can be iteratively obtained as proposed by Kimber and Hollis [27] as follows:

$$L_a^{ro}(0) = \frac{\lambda \hat{x}_a^2}{1-\hat{x}_a} + \hat{x}_a \tag{8}$$

$$\hat{x}_a = \frac{L_a^{ro}(0) + 1 - \sqrt{(L_a^{ro}(0)+1)^2 - 4(1-\lambda)L_a^{ro}(0)}}{2(1-\lambda)} \tag{9}$$

where λ is a constant between 0.4 and 0.6 [27].

It is assumed that the accumulated queue length decreases linearly at a rate which is the difference between the arrival rates corresponding to the degrees of saturation $\hat{x}_a$ and x_a, as it approaches the equilibrium queue length for time slice t. Although Eq. (9) calculates the dissipation time for the accumulated queue length, it underestimates the time taken by the accumulated queue to clear [26]. Using Eq. (9) for each link a uniform delay for undersaturated with accumulated queues can be obtained as follows:

(*i*) if $\tau_a \geq t$ than the uniform component of delay on link a can be calculated as given by Eqs. (4-6).

(ii) if $\tau_a < t$ than the uniform component of delay on link a is calculated by a linear combination of oversaturated and undersaturated conditions as follows:

$$L_a^U = \frac{\mu_a(1-\Lambda_a)C}{2t}\left[\tau_a + \frac{x_a(1-\Lambda_a)(t-\tau_a)}{(1-\Lambda_a x_a)}\right] \tag{10}$$

$$D_a^U = L_a^U \tag{11}$$

$$d_a^U = \frac{D_a^U}{q_a} \tag{12}$$

Following aforementioned statements on link cost function, the upper level objective function can be presented as follows:

Upper level problem:

$$\min TTC(\zeta, \mathbf{q}^*, \boldsymbol{\psi}) = \sum_{a\in A}\left[q_a^*\left(\zeta, \boldsymbol{\psi}\right)\cdot\left(c_a^0 + d_a^U + d_a^{ro}\right)\right] \tag{13}$$

subject to

$$\boldsymbol{\psi}\left(\mathbf{C}, \boldsymbol{\varphi}\right) \in \boldsymbol{\Omega}_0; \begin{cases} C_{\min} \le C \le C_{\max} \\ \varphi_{\min} \le \varphi \le C \\ \sum_{i=1}^{z}\left(\varphi + I\right)_i = C \end{cases} \tag{14}$$

$$q_a^*(\zeta, \boldsymbol{\psi}) \le \mu_a(\boldsymbol{\psi}, s_a) \tag{15}$$

where ζ is O-D demand multiplier, $C_{\min}$ and $C_{\max}$ are possible bounds for cycle time C, φ is stage green time, $\varphi_{\min}$ is the minimum stage green time, I is intergreen time, $\boldsymbol{\psi}$ is vector of signal timings, $\boldsymbol{\Omega}_0$ is feasible region for signal timings, s_a is the saturation flow on link a, and z is the number of stages. Additionally, equilibrium link flow on link a, $q_a^*(\zeta, \boldsymbol{\psi})$ can be obtained by solving SUE problem at the lower level as proposed by Bell and Iida [28] as follows.

Lower level problem:

$$\min_{\mathbf{q}(\boldsymbol{\psi})} F(\mathbf{q}(\boldsymbol{\psi}), \boldsymbol{\psi}) = -\zeta\mathbf{p}^{\mathrm{T}}\mathbf{y}(\mathbf{q}(\boldsymbol{\psi}), \boldsymbol{\psi}) + \mathbf{q}^{\mathrm{T}}\mathbf{c}(\mathbf{q}(\boldsymbol{\psi}), \boldsymbol{\psi}) - \sum_{a\in A}\int_0^{q_a(\boldsymbol{\psi})} c_a(\boldsymbol{\psi}, w)dw \tag{16}$$

subject to

$$\zeta\mathbf{p} = \boldsymbol{\Lambda}\mathbf{h}, \qquad \mathbf{q}\left(\boldsymbol{\psi}\right) = \boldsymbol{\delta}\mathbf{h}, \qquad \mathbf{h} \ge \mathbf{0} \tag{17}$$

where $\mathbf{p}$ is the vector of travel demand, $\mathbf{c}$ and $\mathbf{y}$ represent vectors of link and path travel times for the given vector of link flows $\mathbf{q}(\boldsymbol{\psi})$, respectively, $\mathbf{h}$ is the vector of path flows, $\boldsymbol{\delta}$ represents the link/path incidence matrix where $\delta_{ar}=1$ if link a is on path r, and $\delta_{ar}=0$ otherwise $\left[\delta_{ar};\forall a\in A;\forall r\in R\right]$, and $\boldsymbol{\Lambda}$ is the O-D/path incidence matrix $\left[\Lambda_r;\forall r\in R\right]$.

3. MODEL DEVELOPMENT

As known, *TTC* in a road network can be minimized by optimizing signal timings (i.e. cycle times and stage green times) of isolated intersections. At this point, route choice behavior of road users is considered as a reaction to the changing travel costs based on signal timing adjustments. Therefore, *TTC* minimization problem is formulated as a bi-level programming problem, in which the upper and lower levels represent *TTC* minimization and SUE assignment problems, respectively, as given in Eqs. (13-17). In this study, upper level objective function is modified by transforming the capacity constraint, which is given in Eq. (15), into a penalty function as given in Eq. (18). Note that the second term on the right side of Eq. (18) ensures that the capacity is not violated on any links in the road network.

$$\min TTC(\zeta,\mathbf{q}^*,\boldsymbol{\psi})=\sum_{a\in A}q_a^*(\zeta,\boldsymbol{\psi})\cdot\left(c_a^0+d_a^U+d_a^{ro}\right)+\sum_{a\in A}\sigma\max\left(q_a^*(\zeta,\boldsymbol{\psi})-\mu_a(\boldsymbol{\psi},s_a),0\right) \tag{18}$$

subject to

$$\boldsymbol{\psi}(\mathbf{C},\boldsymbol{\varphi})\in\boldsymbol{\Omega}_0;\quad\begin{cases}C_{\min}\le C\le C_{\max}\\ \varphi_{\min}\le\varphi\le C\\ \sum_{i=1}^{z}(\varphi+I)_i=C\end{cases} \tag{19}$$

where σ is constant weight parameter which is used to include capacity constraint into the objective function given in Eq. (13) as a penalty component. The weight parameter provides a balance between left and right sides of Eq. (18) by reflecting negative effects of links with capacity violation. Considering both non-convexity and the vast search space of the above given optimization problem, the solution is carried out by TRACOM model which is developed based on DE solution framework. The flowchart of TRACOM model is given in Fig. 1.

The stepwise solution procedure of the TRACOM model is given as follows:

Step 1: At this step, the objective function given in Eq. (18), possible bounds for cycle and stage green times, network-related parameters (i.e. free-flow travel times, saturation flows), O-D travel demand, and O-D demand multiplier are initialized. Subsequently, three DE parameters are introduced. These parameters can be explained as follows:

- Population size (*Np*) represents the number of solution namely signal timing vectors in the population pool,

- Mutation factor (F) is used to create a mutant vector,
- Crossover rate (CR) is used to create a trial vector [29].

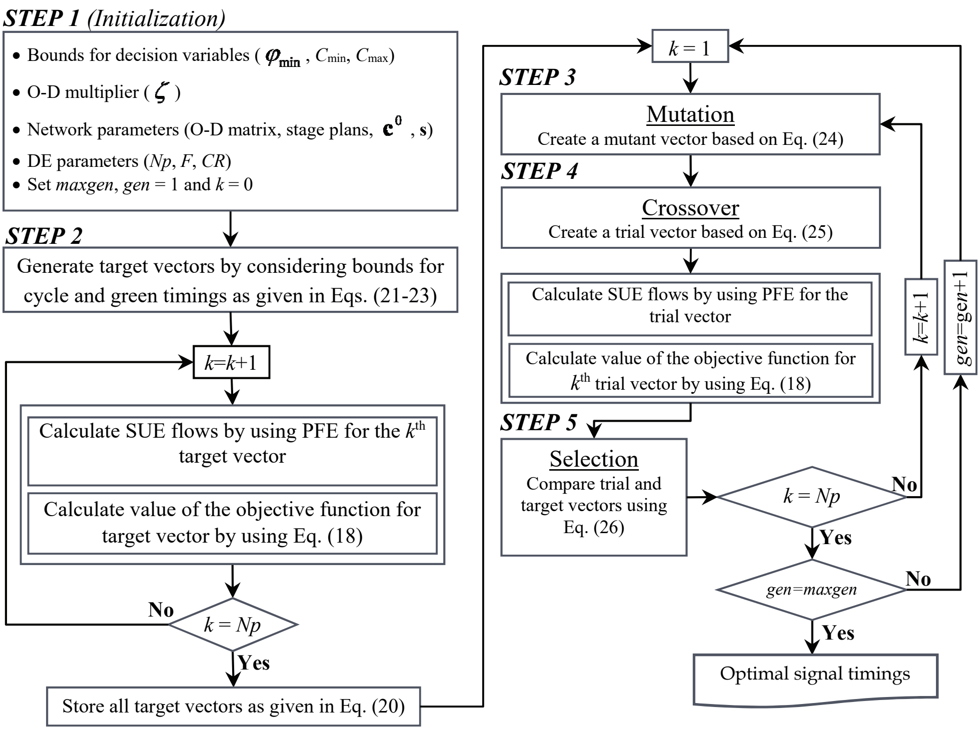


Figure 1 - Flowchart of TRACOM model

Step 2: At this step, initial solution vectors are produced subject to the possible bounds of decision variables which are cycle and green times for each intersection in the road network. Then, TTC values are calculated using Eq. (18) and stored as shown in Eq. (20).

$$\left[\begin{matrix} \overbrace{\begin{matrix} C_1^1 & C_2^1 & \cdots & C_N^1 \\ C_1^2 & C_2^2 & \cdots & C_N^2 \\ \vdots & \vdots & \vdots & \vdots \\ C_1^{Np} & C_2^{Np} & \cdots & C_N^{Np} \end{matrix}}^{\text{Cycle times}} & \overbrace{\begin{matrix} \varphi_{1,1}^1 & \varphi_{1,2}^1 & \cdots & \varphi_{1,z_1}^1 & \varphi_{2,1}^1 & \cdots & \varphi_{N,z_n}^1 \\ \varphi_{1,1}^2 & \varphi_{1,2}^2 & \cdots & \varphi_{1,z_1}^2 & \varphi_{2,1}^2 & \cdots & \varphi_{N,z_n}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \varphi_{1,1}^{Np} & \varphi_{1,2}^{Np} & \cdots & \varphi_{1,z_1}^{Np} & \varphi_{2,1}^{Np} & \cdots & \varphi_{N,z_n}^{Np} \end{matrix}}^{\text{Green times}} \end{matrix}\right] \left[\overbrace{\begin{matrix} TTC\left(\zeta,\mathbf{q}^*,\boldsymbol{\psi}\right)^1 \\ TTC\left(\zeta,\mathbf{q}^*,\boldsymbol{\psi}\right)^2 \\ \vdots \\ TTC\left(\zeta,\mathbf{q}^*,\boldsymbol{\psi}\right)^{Np} \end{matrix}}^{\text{Objective functions}}\right] \qquad (20)$$

where $\varphi_{i,j}$ is j^{th} stage green time of intersection i (i=1,2,…, N and j=1,2,…, z_n), z_n number of stages at n^{th} intersection, and N is the number of intersections. As can be seen in Eq. (20) that the population pool includes initial signal timing vectors and their related *TTC* values as many as *Np*. To generate an initial signal timing vector with TRACOM algorithm, following procedure is executed.

(i). Cycle time values are generated for all intersections in the road network. Following representation is given for the i^{th} intersection:

$$C_i = \text{int}\left[rnd\left(0;1\right)\times\left(C_{\max} - C_{\min}\right) + C_{\min}\right] \tag{21}$$

(ii). Green time values for the stages of all intersections are generated. Eq. (22) illustrates this process for the j^{th} stage of the i^{th} intersection.

$$\varphi_{i,j} = \text{int}\left[rnd\left(0;1\right)\times\left(C_i - \varphi_{\min}\right) + \varphi_{\min}\right] \tag{22}$$

(iii). In order to provide consistency between a cycle time and its components (green and intergreen times), generated green times of each signalized intersection are revised. Eq. (23) illustrates this process for the j^{th} stage of the i^{th} intersection.

$$\varphi_{i,j} = \varphi_{\min} + \frac{\varphi_{i,j}}{\sum_{j=1}^{z_i}\varphi_{i,j}}\left[C_i - z_i \times\left(I + \varphi_{\min}\right)\right] \tag{23}$$

As can be seen in Eq. (18) that the equilibrium link flows, $\mathbf{q}^*$, are required for calculating *TTC* values of signal timing vectors. In TRACOM, the SUE assignment problem is solved using PFE traffic assignment tool. Assuming β is the dispersion parameter, Dell'Orco et al. [17] performed sensitivity analysis on SUE assignment by using PFE and found that for values of β up to 1, the objective function values remain stable, and decrease rapidly for bigger values than 1. This result is reasonable: in fact, the higher the value of β, the more deterministic is the traffic assignment. Therefore, the value of β is selected 1. The pseudo code of the PFE is illustrated in Fig. 2.

Once the SUE assignments are carried out and equilibrium link flows are obtained for each signal timing vector, their corresponding *TTC* values are calculated. Note that as of the end of Step 2, each solution vector in the population is called as a target vector.

Step 3. At this step, a randomly chosen solution vector is mutated by adding a weighted difference of two randomly selected solution vectors. Note that all three vectors must be different from both each other and the target vector, $\mathbf{\Gamma}_i$. A mutant vector, $\boldsymbol{\rho}_i$ is created as follows:

$$\boldsymbol{\rho}_{i,gen} = \mathbf{\Gamma}_{r0,gen} + F\cdot\left(\mathbf{\Gamma}_{r1,gen} - \mathbf{\Gamma}_{r2,gen}\right) \tag{24}$$

where $r0$, $r1$ and $r2$ are indices of randomly chosen solution vectors.

$q_a \leftarrow 0, \forall a \in A$

$c_a \leftarrow c_a(q_a), \forall a \in A$

$m \leftarrow 0$

repeat

$m \leftarrow m+1$

Update link travel costs $c_a \leftarrow \frac{1}{m} c_a(q_a) + (1 - \frac{1}{m}) c_a$

For each path r

Calculate new path costs $y_r \leftarrow \sum_{a \in A} \delta_{ar} c_a(q_a)$

Next r

For each path r

Calculate new path flows $h_r \leftarrow p_k \frac{\exp(-\beta y_r)}{\sum_{r \in \mathbf{R}_k} \exp(-\beta y_r)}$

Next r

For each link a

Calculate new link flows $q_a \leftarrow \sum_{r \in \mathbf{R}_k} \delta_{ar} h_r$

Next a

until no new path and link flows converged

Figure 2 - Pseudo code of the PFE

Step 4. At this step, crossover is applied by choosing each member of the trial vector, $\mathbf{E}_i$ from the target or the mutant vectors with the probabilities of *CR* or 1-*CR*, respectively, as given in Eq. (25).

$$\mathbf{E}_{i,gen} = \varepsilon_{j,i,gen} = \begin{cases} \rho_{j,i,gen} & \text{if} \left(rnd_j(0,1) \le CR \text{ or } j = j_{rnd} \right) \\ \gamma_{j,i,gen} & \text{otherwise} \end{cases} \tag{25}$$

where $\varepsilon_{j,i,gen}$, $\rho_{j,i,gen}$, and $\gamma_{j,i,gen}$ are j^{th} members of i^{th} trial, mutant and target vectors, respectively. The condition of $j = j_{rnd}$ provides that target and trial vectors are definitely different from each other. After determining the members of the trial vector, their related SUE link flows and the *TTC* value is calculated.

Step 5. At the last step, the target vector, $\mathbf{\Gamma}_{i,gen+1}$ for the next generation is selected by comparing the *TTC* values of trial and target vectors as given in Eq. (26).

$$\boldsymbol{\Gamma}_{i,gen+1} = \begin{cases} \mathbf{E}_{i,gen} & \text{if } TTC\left(\mathbf{q}^*\left(\mathbf{E}_{i,gen}\right), \mathbf{E}_{i,gen}\right) \leq u\left(\mathbf{q}^*\left(\boldsymbol{\Gamma}_{i,gen}\right), \boldsymbol{\Gamma}_{i,gen}\right) \\ \boldsymbol{\Gamma}_{i,gen} & \text{otherwise} \end{cases} \tag{26}$$

The DE process is repeated until the maximum number of generations, *maxgen*, is reached. In order to provide a more explanatory illustration, pseudo code of TRACOM model is given in Fig. 3.

---

for $\zeta \leftarrow 1$ to $\zeta_{\max}$ step α **do** (α is defined as step size for O-D multiplier, ζ)
    for $k \leftarrow 1$ to Np **do**
        for $i \leftarrow 1$ to N **do**
            Generate a cycle time C_i for i^{th} signalized intersection randomly from $\in \{ C_{\min}, \ldots, C_{\max} \}$
            for $j \leftarrow 1$ to z_i **do**
            Generate a green split $\varphi_{i,j}$ for j^{th} stage of i^{th} signalized intersection randomly from $\in \{\varphi_{\min}, \ldots, C_i\}$
            Revise stage green times of i^{th} signalized intersection based on Eq. (23)
        Run PFE to calculate SUE flows $\mathbf{q}^*$ for the k^{th} initial signal timing vector
        Calculate total travel cost for k^{th} target (initial signal timing) vector and current O-D matrix multiplier ζ
    for *gen* $\leftarrow$ 1 to *maxgen* **do**
        for $k \leftarrow 1$ to Np **do**
            Perform mutation to the k^{th} solution vector to create a mutant vector based on Eq. (24)
            for $v \leftarrow 1$ to Nd **do** (*Nd* is defined as number of decision variables)
                Perform crossover to obtain the v^{th} decision variable of k^{th} trial vector based on Eq. (25)
            Revise intersection cycle times considering their upper and lower bounds $\{ C_{\min}, \ldots, C_{\max} \}$
            Revise stage green times based on Eq. (23)
            Run PFE to calculate SUE flows $\mathbf{q}^*$ for the trial vector
            Calculate total travel cost for trial vector and current O-D matrix multiplier ζ
            Replace k^{th} target vector in the population with the k^{th} trial vector in case it produces lower total travel cost.
    print optimal/near optimal signal timings and total travel cost for current O-D matrix multiplier ζ

Figure 3 - Pseudo code of TRACOM

4. NUMERICAL APPLICATION

Medium-sized signalized road network firstly presented by Allsop and Charlesworth [30] is chosen for numerical application of TRACOM model. Additionally, by using this benchmark road network we aimed to give readers a chance to compare some results drawn by previous studies in the literature since it has widely been used in most of traffic related studies. Allsop and Charlesworth's network has 23 links and 20 signal setting variables at six junctions by taking each one as isolated into account. The representation of the road network and its stage plans are given in Figs. 4 and 5 while the network data and O-D travel demand matrix are presented in Tables 2 and 3, respectively.

Table 2 - Network data

Intersection number	Link number	Free-flow travel time (c_a^0)	Saturation flow (s_a)	Intersection number	Link number	Free-flow travel time (c_a^0)	Saturation flow (s_a)
1	1	1	2000	4	5	20	1800
	2	1	1600		6	20	1850
	16	10	2900		10	10	2200
	19	10	1500		11	1	2000
					12	1	1800
					13	1	2200
2	3	10	3200	5	8	15	1850
	15	15	2600		9	15	1700
	23	15	3200		17	10	1700
					21	15	3200
3	4	15	3200	6	7	10	1800
	14	20	3200		18	15	1700
	20	1	2800		22	1	3600

Table 3 - O-D travel demand matrix (veh/h)

O-D	A	B	D	E	F
A	--	250	700	30	200
C	40	20	200	130	900
D	400	250	--	0	100
E	300	130	0	--	20
G	550	450	170	60	20

Constraints for each signal timing variable used in TRACOM are set as given below:

$$\boldsymbol{\psi}\left(\mathbf{C},\boldsymbol{\varphi}\right)\in\boldsymbol{\Omega}_0;\ \begin{cases}36\le C\le 120\\ 7\le\varphi\le C\end{cases} \tag{27}$$

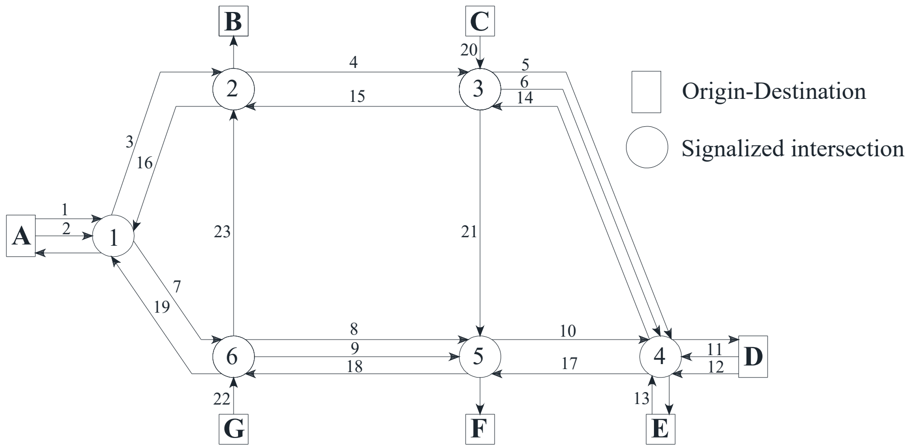


Figure.4 - Allsop & Charlesworth's network

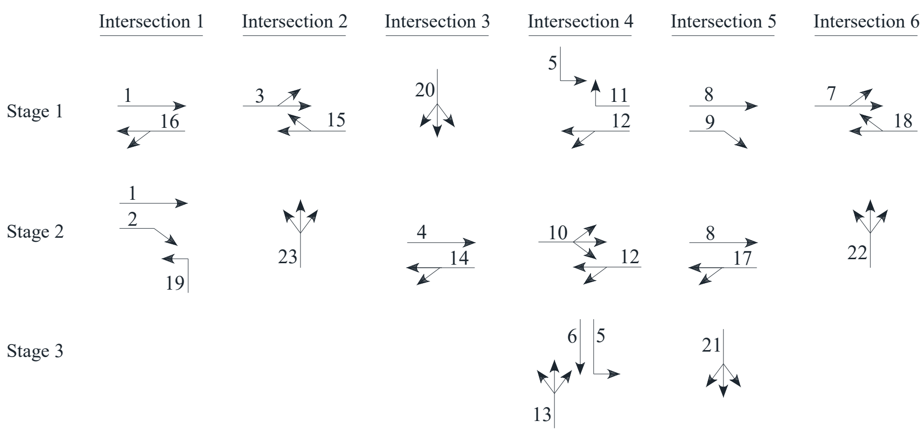


Figure 5 - Stage plans

Considering that the values of DE parameters play an important role on the performance of TRACOM model, a sensitivity analysis has been performed based on the recommended ranges for DE parameters by Storn and Price [31]. These ranges are assumed to be [0.5, 1.0] and [0.8, 1.0] for mutation factor and crossover rate, respectively. For this purpose, 30 different cases are created and each case is solved 10 times with different random seeds for base O-D travel demand matrix. Note that the intergreen time, *I*, between stages was set to 5 seconds, population size, *Np*, was set to 30, and maximum number of generations, *maxgen*, was set to 1000 during the sensitivity analysis. The best *TTC* values after 10 solutions for each case are given in Table 4.

Table 4 - TTC values resulting from the sensitivity analysis (veh-h)

CR \ F	0.50	0.60	0.70	0.80	0.90	1.00
0.80	148.02	147.94	147.49	147.28	148.53	149.66
0.85	148.02	147.94	147.56	147.54	148.91	149.66
0.90	148.55	148.91	148.14	149.49	149.95	149.66
0.95	148.97	148.69	148.69	149.85	149.49	149.78
1.00	148.97	148.97	148.79	149.04	150.23	150.11

As can be seen in Table 4 that the minimum TTC value was obtained as 147.28 veh-h for the case with both *F* and *CR* are 0.80. The computational time for complete run of TRACOM resulted in 2.61 hours. It means that each generation takes about 9.4 seconds of CPU. TRACOM has been executed in MATLAB programming and performed on PC with Intel Core i7 2.10 GHz, RAM 8 GB. Based on the results obtained with the sensitivity analysis, the following user-specified DE parameters were used during the analyses for Allsop and Charlesworth's network: *CR*=0.8, *F*=0.8, *Np*=30, and *maxgen*=1000. The upper bound for the O-D multiplier, ζ_{max} and step size, α are set to 1.30 and 0.02, respectively. *TTC*, ζ and their changes are given in Table 5 after applying the TRACOM model.

Table 5 - Evolution of the change of ζ and TTC

i	O-D demand multiplier ζ_i	Total travel cost TTC_i	Change (%) $\Delta\zeta = \frac{\zeta_i - \zeta_{i-1}}{\zeta_{i-1}}$	Change (%) $\Delta TTC = \frac{TTC_i - TTC_{i-1}}{TTC_{i-1}}$
1	1.00	147.28		
2	1.02	153.80	2.00	4.43
3	1.04	163.00	1.96	5.98
4	1.06	170.76	1.92	4.76
5	1.08	179.10	1.89	4.88
6	1.10	186.53	1.85	4.15
7	1.12	197.50	1.82	5.88
8	1.14	208.80	1.79	5.72
9	1.16	218.39	1.75	4.59
10	1.18	243.15	1.72	11.34

As can be seen in Table 5 that network *TTC* value is increased up to 65% while the O-D demand matrix is increased up to 18%. On the other hand, when ζ equals 1.20, degree of saturation of at least one link exceeds 100% that means the travel demand can be increased only up to 18% by optimizing intersection signal timings. It can also be seen in Table 5 that, changes of *TTC* vary between 4%-6% at each step *i* during the O-D multiplier increases from 1.00 to 1.16. On the other hand, change of *TTC* is about 11% when ζ equals 1.18 that can also be seen as a sudden spike in Fig. 6. For this reason, the critical value of O-D multiplier can be selected as 1.16.

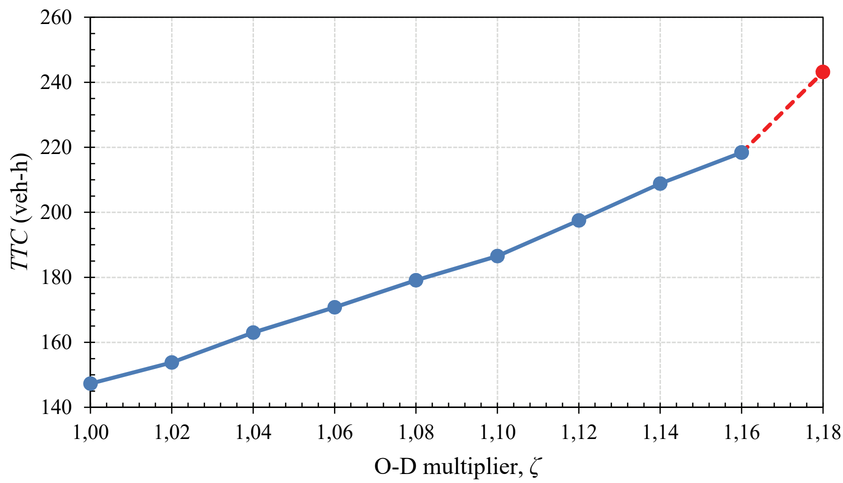


Figure 6 - Relationship between TTC and O-D multiplier ζ

In order to illustrate the convergence behavior of the TRACOM model, Fig. 7 is presented for ζ=1.16 since this value is selected as a critical point for capacity enhancement on the road network to cope with increasing travel demand. As can be seen in Fig. 7, the TRACOM algorithm begins to search optimal signal timings in the feasible search space and finds the initial value of *TTC* at about 435 veh-h. After 10th generation, the TRACOM starts to ignore worse solution vectors and seriously improves the value of *TTC* about 36%. After continuing to improve the objective function, the TRACOM finds the final value of *TTC* as about 218 veh-h. That is, the total improvement rate on the objective function value is about 100% after the1000th generation.

For further information, SUE link flows and corresponding degree of saturations found for ζ=1.16 are given in Table 6. As shown, there is no links exceeding their capacities in the road network since degree of saturations of links are less than 100%. It should be pointed out that the degree of saturations of links numbered 10, 13, 20, 21 and 22 are higher than 90% because SUE flows of those links approach their capacities. Additionally, optimal signal timings by applying TRACOM algorithm are presented in Table 7. By providing cycle time constraints, the lowest cycle time is found as 55 seconds for the second intersection while the highest cycle time is 111 sec for the fifth intersection.

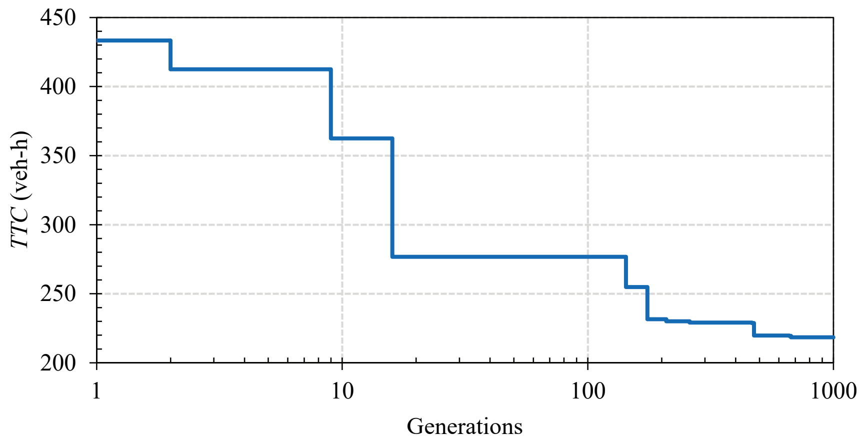


Figure 7 - Convergence graph of TRACOM for ζ =1.16

Table 6 - SUE link flows and corresponding degree of saturations for ζ =1.16

Link number	SUE link flows (veh/h)	Degree of saturation (%)	Link number	SUE link flows (veh/h)	Degree of saturation (%)
1	833	47	13	522	91
2	536	61	14	914	82
3	833	65	15	915	88
4	687	61	16	768	78
5	740	72	17	478	87
6	203	42	18	407	52
7	536	65	19	728	88
8	554	68	20	1496	98
9	128	64	21	1240	92
10	554	93	22	1450	93
11	578	92	23	984	73
12	292	28			

Table 7 - Optimum signal timings for ζ =1.16

Cycle time C (s)	Intersection number i	Duration of stages (s)		
		$\varphi_{i,1}$	$\varphi_{i,2}$	$\varphi_{i,3}$
94	1	32	52	---
55	2	22	23	---
97	3	53	34	---
96	4	30	26	25
111	5	13	36	47
92	6	42	40	---

5. DISCUSSION AND CONCLUSION

The urban population in developing countries is mostly projected to continue to increase. This is likely to mean more drivers resulting in more traffic day by day. On one hand road users expect reliable journeys, on the other hand local authorities focus on maintaining the traffic operations successfully by implementing traffic control strategies, clearing up the incidents quickly, keeping the lanes open. When these tools remain incapable of overcoming traffic congestions, authorities take capacity enhancement countermeasures such as increasing link capacities, lane additions or applying grade-separated junctions. However, such actions bring high investment costs. Therefore, it is crucial to determine the right time for such physical improvements to avoid both premature investments considering limited resources, and late investments to maintain traffic flow properly.

In this study, effectiveness of network wide signal timing optimization on total travel cost in case of increasing travel demand is investigated. For this purpose, a bi-level programming model TRACOM is developed. At the upper level, network TTC, which is a function of link traffic volumes, free flow travel times, uniform and random plus over saturation components of delay, is minimized. At the lower level of the proposed model, drivers' route choice behaviors are taken into consideration in SUE context. The TRACOM model, which is based on DE optimization method, is applied to Allsop and Charlesworths' test network to evaluate its effectiveness. In the case of Allsop and Charlesworth's road network, the traffic agency can manage the network by optimizing signal timings until the O-D travel demand increases 18%. However, the TTC shows a sudden spike while the value of O-D multiplier is between 1.16 and 1.18 although it shows an approximate linear increase while the travel demand is increased up to 16%. Thus, any capacity enhancement countermeasure before or after the 16% increase in travel demand would be a premature or late investment, respectively.

Capacity enhancement represents improvement of an existing road network by investing in new transport construction which plays an important role on the performance of the road network. Timely and efficient investments in roads provide economic and social benefits to the emerging economies. On the other hand, premature investments waste limited resources and it may hinder the realization of more important services. In this context, local authorities who are responsible for urban road network management can benefit from TRACOM model to make their investment decisions timely based on projected travel demand data provided in transportation master plans.

The TRACOM model considers networks including signalized intersections. Since the use of different types of intersections (i.e. stop controlled or roundabout) will clearly affect the results of TRACOM model, this issue will be taken into account in future studies. Moreover, different multipliers for each O-D pair can be used rather than a common O-D matrix multiplier. The effects of signal coordination will also be taken into consideration.

References

[1] Webster F. V., Traffic signal settings. Road Research Technical Paper, 39. HMSO, London, 1958.

[2] Allsop, R. E., Delay-minimizing settings for fixed-time traffic signals at a single road junction. Journal of the Institute of Mathematics and its Applications, 8, 164–185, 1971.

[3] Allsop, R. E., Estimating the traffic capacity of a signalized road junction. Transportation Research, 6, 245–255, 1972.

[4] Wong, S. C., Derivatives of performance index for the traffic model from TRANSYT. Transportation Research Part B, 29, 303–327, 1995.

[5] Heydecker, B. G., A decomposed approach for signal optimisation in road networks. Transportation Research Part B, 30, 2, 99–114, 1996.

[6] Wong, S. C., Group-based optimisation of signal timings using the TRANSYT traffic model. Transportation Research Part B, 30, 217–244, 1996.

[7] Wong, S. C., Group-based optimisation of signal timings using parallel computing. Transportation Research Part C, 5, 123–139, 1997.

[8] Wong, S. C., Wong, W. T., Leung, C. M., and Tong, C. O., Group-based optimization of a time-dependent TRANSYT traffic model for area traffic control. Transportation Research Part B, 36, 4, 291–312, 2002.

[9] Girianna, M., and Benekohal, R. F., Application of genetic algorithms to generate optimum signal coordination for congested networks. Proceedings of the Seventh International Conference on Applications of Advanced Technology in Transportation, Cambridge, MA, United States, 762–769, 2002.

[10] Ceylan, H., and Bell, M. G. H., Traffic signal timing optimisation based on genetic algorithm approach, including drivers' routing. Transportation Research Part B, 38, 4, 329–342, 2004.

[11] Ceylan, H., Developing combined genetic algorithm hill-climbing optimization method for area traffic control. Journal of Transportation Engineering, 132, 8, 663–671, 2006.

[12] Chen, J., and Xu, L., Road-junction traffic signal timing optimization by an adaptive particle swarm algorithm. 9th International Conference on Control, Automation, Robotics and Vision, Singapore, 1-7, 2006.

[13] Dan, C., and Xiaohong, G., Study on intelligent control of traffic signal of urban area and microscopic simulation. Proceedings of the Eighth International Conference of Chinese Logistics and Transportation Professionals, Logistics: The Emerging Frontiers of Transportation and Development in China, Chengdu, China, 4597–4604, 2008.

[14] Li, Z., Modeling arterial signal optimization with enhanced cell transmission formulations. Journal of Transportation Engineering, 13, 7, 445–454, 2011.

[15] Liu, Y., and Chang, G.-L., An arterial signal optimization model for intersections experiencing queue spillback and lane blockage. Transportation Research Part C, 19, 130–144, 2011.

[16] Ceylan, H., and Ceylan, H., A Hybrid Harmony Search and TRANSYT hill climbing algorithm for signalized stochastic equilibrium transportation networks. Transportation Research Part C, 25, 152–167, 2012.

[17] Dell'Orco, M., Baskan, O., and Marinelli, M., A Harmony Search algorithm approach for optimizing traffic signal timings. Promet Traffic & Transportation, 25, 4, 349–358, 2013.

[18] Dell'Orco, M., Baskan, O., and Marinelli, M., Artificial bee colony-based algorithm for optimising traffic signal timings. Soft Computing in Industrial Applications, Advances in Intelligent Systems and Computing, 223, Eds: Snášel, V., Krömer, P., Köppen, M., Schaefer, G., Springer, Berlin/Heidelberg, 327–337, 2014.

[19] Ozan, C., Baskan, O., Haldenbilen, S., and Ceylan, H., A Modified Reinforcement Learning Algorithm for Solving Coordinated Signalized Networks. Transportation Research Part C, 54, 40–55, 2015.

[20] Christofa, E., Ampountolas, K., and Skabardonis, A., Arterial traffic signal optimization: A person-based approach. Transportation Research Part C, 66, 27–47, 2016.

[21] Srivastava, S., and Sahana, S. K., Nested hybrid evolutionary model for traffic signal optimization. Applied Intelligence, 46, 113–123, 2017.

[22] Abdul Aziz, H. M., Zhu, F., and Ukkusuri, S. V., Learning based traffic signal control algorithms with neighborhood information sharing: An application for sustainable mobility. Journal of Intelligent Transportation Systems, 22, 1, 40-52, 2018.

[23] Bell, M. G. H., Cassir, C., Grosso, S., and Clement, S. J., Path Flow Estimation in Traffic System Management. IFAC Proceedings Volumes, 30, 8, 1247-1252, 1997.

[24] Vincent, R. A., Mitchell, A. I., and Robertson, D. I., User guide to TRANSYT version 8. TRRL Report, LR888. Transport and Road Research Laboratory, Crowthorne, 1980.

[25] Ceylan, H., A genetic algorithm approach to the equilibrium network design problem. PhD Thesis, University of Newcastle upon Tyne, UK, 2002.

[26] Chiou, S-W., Optimisation of Area Traffic Control for Equilibrium Network Flows. PhD Thesis, University College London, 1998.

[27] Kimber, R. M., and Hollis, E. M., Traffic queues and delays at road junctions. TRRL report, LR909. Transportation and Road Research Laboratory, Crowthorne, 1979.

[28] Bell, M. G. H., and Iida, Y., Transportation Network Analysis. John Wiley and Sons, Chichester, UK, 1997.

[29] Storn, R., and Price, K., Differential Evolution: A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces. ICSI, Technical Report, TR-95-012, USA, 1995.

[30] Allsop, R. E., and Charlesworth, J. A., Traffic in a signal-controlled road network: an example of different signal timings including different routings. Traffic Engineering Control, 18, 5, 262-264, 1977.

[31] Storn, R., and Price, K., Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. Journal of Global Optimization, 11, 4, 341-359, 1997.

Teknik Dergi, 2020 9919-9940, Paper 573

Identifying Factors that Contribute to Severity of Construction Injuries using Logistic Regression Model

Özge AKBOĞA KALE[1]
Selim BARADAN[2]

ABSTRACT

Majority of research in occupational safety and health area lean towards describing accidents with the aid of surveys and descriptive statistics, instead of using inferential statistical techniques. Therefore, an extensive archival study was performed in cooperation with Social Security Institute of Turkey, which included examination and reorganization of more than 2000 accident report forms to create a categorically identified data set, incorporating "Injury Severity Score" concept, followed by various statistical analysis techniques (univariate frequency, cross tabulation and binary logistic regression). As a result, a model was developed to identify the factors that contribute to severity. The findings of the analyses showed that four of the independent variables (work experience, accident type, unsafe condition and unsafe act) have statistically significant influence on workplace injury severity.

Keywords: Occupational safety and health, logistic regression analysis, injury severity score, construction accidents, data mining.

1. INTRODUCTION

Research efforts in occupational safety and health field have mainly concentrated on determining the causes of occupational injuries and illnesses and thus help discover new strategies to reduce or eliminate them. Majority of research studies in the past used surveys to collect information from employers and workers [1-3], while several studies preferred to use descriptive statistics to summarize and interpret occupational injury and illness data [4-12]. Even though conducting surveys and using descriptive statistics may be helpful, such as describing how the accident occurred and who were involved; using them alone is not sufficient in determining the factors that contribute most to the unfortunate incident. Therefore, using more advanced statistics techniques, such as logistic regression, to support

Note:
- This paper has been received on October 15, 2018 and accepted for publication by the Editorial Board on January 28, 2019.
- Discussions on this paper will be accepted by May 31, 2020.

- https://dx.doi.org/10.18400/tekderg.470633

1 İzmir Demokrasi University, Department of Civil Engineering, Izmir, Turkey - ozge.akbogakale@idu.edu.tr - https://orcid.org/0000-0002-3848-0578

2 Ege University, Department of Civil Engineering, Izmir, Turkey- selim.baradan@ege.edu.tr - https://orcid.org/0000-0002-9172-8552

findings coming from surveys and descriptive statistics could be invaluable in accident causation studies.

Logistic regression, which is a commonly used inferential statistics technique in medical and social science to develop statistical models and make predictions [13-17], is rarely utilized in occupational safety and health research [18-20], most likely, this study aims to fill this gap in occupational safety and health field by integrating inferential statistics, namely logistic regression technique, to develop a statistical model in identifying the factors that contribute most to the severity of occupational injuries. The study also benefited from traditional descriptive statistics techniques to analyze frequency distribution, which are discussed in the following sections.

Severity is particularly an important topic to focus on. Risk is typically defined as the combination of two components: the probability (frequency) of a defined hazard and the consequences of its occurrence (severity). In occupational safety and health field, severity is more difficult to estimate when compared to frequency. Frequency can be merely calculated by finding the ratio between number of incidents and total number of workers or loss of time due to injury and total work hours. On the other hand, severity is the potential loss when an event occurs and the loss may be expressed in human terms, such as loss of life, serious injury, serious illness, number of losses and so forth [21]. Therefore, efforts were made in this study to find methods to properly quantify and analyze severity component of occupational safety and health risk.

The reason for the drought of advanced statistical studies in occupational safety and health research is most likely due to available public data being ambiguous and insufficient. Similarly, a detailed database was paramount for the intended study to proceed. Thus, cooperation was established with Social Security Institution (SSI), which collects injury statistics in Turkey, to access their database for archival study. Turkish construction industry was used as the sample data set, since construction is the leading industry in fatal injuries (35.3%) in Turkey, and still remains a major problem with ever increasing construction projects and catastrophic incidents along with them [22-24], Consequently, this study also aims to depict an accurate safety portrait of Turkish construction industry, taking into account that there are not many statistical studies due to improper recordkeeping system and lack of immediate data in construction.

The objectives of the research study addressed in this paper can be summarized as: (a) identifying the factors affecting the injury severity score of construction injuries; (b) investigating the frequency distributions of these factors; (c) examining the relationships between the factors; and (d) developing a predictive statistical model for construction injuries in Turkey.

2. MATERIALS AND METHODS

The research study summarized in this paper consisted of two parts: The first part was mainly the creation of database to be used in statistical analyses. Physical reports of more than two thousand construction accidents that resulted in fatal or nonfatal injury were collected from the offices of Social Security Institution in three different cities (İstanbul, Ankara, İzmir), since there was not an available online database. Then the raw data in the reports were organized to generate a new database to be used for statistical analyses. The database included 16 research variables, where each variable had several categories. Injury severity score was selected as the dependent variable, while the others were assumed as independent variables.

The second part of the study was applying a methodology that included performing three different statistical analysis techniques on the database. First, descriptive univariate frequency analysis was performed to establish data demographics and observe the distribution of selected variables with using frequency tables. Then in the next step, cross-tabulation, a bivariate analysis method, was used to examine the relationships between each independent variable and the dependent variable. The final statistical analysis in the study was applying binary logistic regression on the identified independent variables to come up with a model that predicts the injury severity. The flow chart of the methodology is displayed in Figure 1.

The most challenging parts of the methodology were to quantify injury severity for logistic regression, where many factors affect the outcome, and how to integrate fatal and nonfatal injuries. Therefore, a literature review was conducted to find a universally accepted method to quantify injury severity. Even though the review failed to find such a method used in construction safety and health area, a rating system named "Injury Severity Score (ISS)", which is predominantly seen in medicine and traffic safety research studies, was adopted for this study. Using ISS to quantify severity of construction injuries enabled ranking and making comparisons between each injury case. Furthermore, ISS was a great fit for logistic regression analysis part of the study due to ISS having a nonlinear structure.

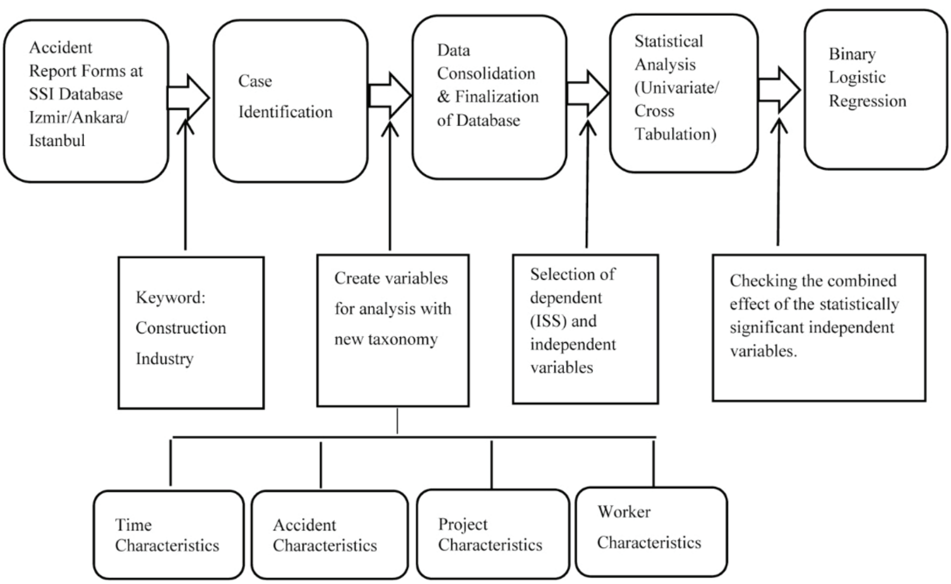


Figure 1 - Flow chart of the methodology

3. DATA ACQUISITION AND MINING

Finding occupational injury data is a significant challenge in Turkey, even though employers are responsible for reporting worksite injuries and illnesses to Social Security Institution (SSI) as mandated by law (Article 14) [25]. SSI shares annual summaries of these reports on their website [22]. However, it isn't possible to extract detailed information from the available online data. Therefore, a special work permit was obtained to be able to study injury reports stored in SSI's physical archives in three biggest cities of Turkey (Ankara, Istanbul, and Izmir).

Following this arrangement, a systematic archival study was carried out. Since, work-related injury reports lack an industry or injury based categorization system, report forms associated with construction injuries had to be handpicked among 50,000 work-related accident cases. This data collection effort covered only the cases reported between 2010 and 2012. During data collection process there have been no important changes in the way that accident data were collected or in compensation practices. The classification system of accidents and occupations remained unchanged as well.

The selection process included identifying fatal and nonfatal injuries that only occurred in the construction sites and production areas, such as ready-mix concrete, prefabricated structural concrete and steel production. Accident cases in other industries were not taken into consideration. Overall 2249 construction related cases were found in the database (1062 cases from Izmir, 732 cases from Istanbul, and 655 cases from Ankara) of those, 425 records were disregarded due to inadequate or missing information. Finally, a total of 2024 cases were selected for analyses.

The SSI workplace accident report forms already contain some of the information about the variables within the defined characteristics, such as injured worker characteristics, nature of injury, injured body part and equipment involved. However, existing information wasn't sufficient to perform the intended analysis in this study. Particularly, there were too many inconsistent answers due to the nature of open ended questions contained in the form. Therefore, a new taxonomy was created to be used for descriptive statistical analysis. During this process, a thorough screening of collected data was performed to identify missing data, consolidate some of the categories and include additional variables.

The significance of the pattern of missing data in a data set is indicated by Tabachnik and Fidell (2013) [26]. Deleting the variables or dropping the cases with missing data from the data set are two ways to deal with missing data. Dismissing or dropping the case from the data set is the first alternative if a case is missing too many data. However, just deleting the variable is suggested if only certain variable information is missing for too many cases. The data set of this study had only one problematic variable in terms of missing data; worker's education background. Apart from this, there were also a few unknown variables in several cases. However, all the cases were kept to acknowledge their presence in the database. Only the missing data from certain variables were neglected and weren't included in the analysis.

Additional categorical variables were also required to be added to the new database by using the information extracted from open-ended questions. Kass (1980) also suggests merging some levels in order to reach a meaningful conclusion in studies that involve data mining [27]. Therefore, collapsing levels technique was applied to variables with high number of categories but low number of observation counts.

Finally, taxonomy study resulted in 15 independent research variables. They were grouped under four different categories (time, project, accident, and worker characteristics) according to their relevance with their characteristics, as seen in Tables 1-4. Also, Injury Severity Score was selected as the dependent variable required for bivariate analysis section of the study, which will be addressed in the following section. Independent variables selected for this research were, in most part, consistent with those used in past research [28-33]. More information about the categories and variables listed under them are given below.

Time characteristics describe when the accident happened and include three variables: month of the year, day of the week and hour of the day.

Table 1 - Distribution of Time Characteristics

Variable	Description	Frequency		Variable	Description	Frequency	
		Count	Percentage (%)			Count	Percentage (%)
Month of accident	January	205	10.1	**Hour of Accident**	$07.00<t\leq 08.00$	19	1.1
	February	207	10.2		$08.00<t\leq 09.00$	152	8.5
	March	208	10.3		$09.00<t\leq 10.00$	216	12.0
	Aprıl	181	8.9		$10.00<t\leq 11.00$	225	12.5
	May	171	8.4		$11.00<t\leq 12.00$	158	8.8
	June	153	7.6		$12.00<t\leq 13.00$	72	4.0
	July	166	8.2		$13.00<t\leq 14.00$	174	9.7
	August	117	5.8		$14.00<t\leq 15.00$	189	10.5
	September	132	6.5		$15.00<t\leq 16.00$	196	10.9
	October	182	9.0		$16.00<t\leq 17.00$	157	8.8
	November	112	5.5		$17.00<t\leq 18.00$	81	4.5
	December	190	9.4		$18.00<t\leq 19.00$	40	2.2
Day of Accident	Monday	349	17.2		$19.00<t\leq 20.00$	23	1.3
	Tuesday	296	14.6		$20.00<t\leq 07.00$	92	5.1
	Wednesday	320	15.8				
	Thursday	313	15.5				
	Friday	305	15.1				
	Saturday	253	12.5				
	Sunday	188	9.3				

Project characteristics provide information about the construction project that the worker was involved in (type and end use) at the time of incident. Construction sites, different in shape and size, are unique dynamic environments. Therefore, project characteristics help to understand and classify the construction environment where accidents mostly occur. Project type variable comprises of 3 selections: new project or addition, manufacturing of construction materials, and repair/maintenance/renovation of existing structures. Project end use variable addresses the usage purpose of the finalized project (residential, institutional and commercial, industrial, infrastructure and construction materials).

Table 2 - Distribution of Project Characteristics

Variable	Description	Frequency		Variable	Description	Frequency	
		Count	Percentage (%)			Count	Valid Percentage (%)
Project Type	New project or new addition	1710	84.5	Project End Use	Residential	1334	66.3
	Manufacturing of construction materials	255	12.6		Construction Materials	255	12.7
	Repair/Maintenance/ Renovation	59	2.9		Institutional and Commercial	243	12.1
					Infrastructure / Heavy Construction	167	8.3
					Industrial	13	0.6

Accident characteristics variables (type of injury, nature of injury, injured body part, unsafe act and condition) reveal plenty of information regarding the incident and injury; in other words, they describe the accident. Categorical data under “type of injury” variable were created based on the International Classification of Diseases (ICD) codes [34], while nature of injury and injured body part information were directly taken from workplace accident reports. Unsafe acts and conditions variables were deducted from the open ended questions in the workplace accident report forms based on the ILO accident cause theory and Hill’s definitions [35-36]. Heinrich's Domino Theory states that accidents result from a chain of sequential events, metaphorically like a line of dominoes falling over. According to Herbert W. Heinrich; all incidents directly relate to unsafe conditions and acts and removing a key factor (an unsafe condition or an unsafe act) prevents the start of the chain reaction [37].

Worker characteristics variables aim to investigate the personal (Age, education) and professional (work experience, assigned task, construction trade) background of injured worker. Data categories were mostly determined in compliance with the classification used in SSI annuals.

Table 3 - Distribution of Accident Characteristics

Variable	Description	Frequency		Variable	Description	Frequency	
		Count	Percentage (%)			Count	Percentage (%)
Injured Body part	Upper extremity	623	32.7	Nature of Injury	Superficial wound	496	26.2
	Lower extremity	444	23.3		Bruise	379	20.0
	head	359	18.9		Fracture/Crack	262	13.8
	Multi-injury	218	11.5		Cut	235	12.4
	Whole body	118	6.2		Skeletal and muscular system disorders	132	7.0
	Internal organs	81	4.3		Death	99	5.2
	Back	50	2.6		Foreign object in the eye	91	4.8
	Neck	10	0.5		Strain	75	4.0
Unsafe condition	Insufficient/Lack of Written Work Procedures	640	32.1		Trauma and Internal injury	67	3.5
	Poor Housekeeping	484	24.3		Electrocution	26	1.4
	No collective protection systems	387	19.4		Burns (heat or chemical)	24	1.2
	No PPE provided	293	14.7		Others	10	0.5
	Faulty tool/equipment/machinery	191	9.5	Type of injury	Falls	734	36.4
Unsafe Act	No PPE Usage	678	34.0		Struck by falling object	604	30.0
	Unsafe work practices	604	30.3		Bite/Sting/Scratch	283	14.1
	Position Inappropriate for Task	312	15.6		Caught In or Between	260	12.9
	Unsafe act by a third party	222	11.1		Others	52	2.6
	Defective/Inappropriate Equipment In Use	179	9.0		Electric shock	29	1.4
					Vehicle accident	24	1.2
					Fire / Explosion	18	0.9
					Heat exposure	10	0.5

Table 4 - Distribution of Worker Characteristics

Variable	Description	Frequency		Variable	Description	Frequency	
		Count	Percentage (%)			Count	Percentage (%)
Worker's age	14-15	1	0.1	**Assigned task**	Formwork	494	25.3
	16-18	36	2.0		Material handling	342	17.5
	19-24	373	20.3		Plaster/Paint	174	8.9
	25-29	362	19.7		Assembly/disassembly	168	8.6
	30-34	328	17.9		Ironwork	155	7.9
	35-39	245	13.4		Concrete handling	118	6.0
	40-44	211	11.5		Maintenance / Repair	98	5.0
	45-49	176	9.6		Other services	91	4.7
	50-54	65	3.5		Commuting	58	3.0
	55-59	32	1.7		Field Inspection	54	2.8
	60-64	3	0.2		Woodwork	44	2.3
	65+	1	0.1		Welding	42	2.1
Work experience in project	Beginning at same day	167	8.6		Work break	35	1.8
	1. Week	170	8.7		Installation	23	1.2
	2-4 Weeks	455	23.4		Excavation work	20	1.0
	1-3 Months	551	28.3		Sheet metal work	18	0.9
	3-6 Months	289	14.8		Marble and tile setting	16	0.8
	6-12 Months	166	8.5		Asphalt paving	5	0.2
	12-24 Months	40	2.1	**Construction trade**	Unskilled worker	563	28.1
	>24 Months	110	5.6		Form worker	512	25.5
Educational Background	Literate	93	8.8		Ironworker	181	9.0
	Illiterate	21	2.0		Assembly/Installation/Maintenance	162	8.1
	Elementary school	567	53.7		Plasterer/Painter	159	7.9
	Middle school	317	30.0		Heavy equipment operator	112	5.6
	High school	29	2.7		Foreman	97	4.8
	Vocational high school	11	1.0		Welder	64	3.2
	Undergraduate	16	1.5		Plumber	40	2.0
	Graduate	1	0.1		Technician	36	1.8
					Installation	26	1.3
					Tile setter	17	0.8
					Security guard	14	0.7
					Electrician	13	0.6
					Carpenter	10	0.4

4. DATA ANALYSIS

4.1. Injury Severity Score (ISS)

Appropriate classification of injuries by type and severity is fundamental to the study of injury [38]. Even within a single community, groups of injured persons differ as to the nature and severity of their injuries. The difficulty of adjusting for such variation has hampered scientific study of injured persons. Nevertheless, it is essential to take differences in severity of injury into account when comparing the morbidity (the state of being diseased or unhealthy) and mortality (the number of people who died) of various groups for purposes of evaluating their emergency and subsequent care [39].

Quantifying injury severity has been a challenge in this study as well. A literature review was conducted to find a universally accepted method to quantify injury severity. The Injury Severity Score (ISS), a method predominantly seen in medicine and traffic safety research studies, was found. ISS is virtually the only anatomical scoring system in use, and correlates linearly with mortality, morbidity, hospital stay and other measures of severity [40]. Therefore, this study adapted ISS rating system to quantify severity of construction injuries, and to rank and make comparisons between each injury case. It was also deemed a great fit for logistic regression analysis part of the study due to its nonlinear structure.

Table 5 - Classification of ISS

ISS	Consequence	Symptom	Count	Percentage (%)
≤ 3	First aid was not needed. No time away from work.	Strain, skeletal and muscular system disorders	441	23.3
4 ≤ ISS ≤ 8	First aid was given but no medical intervention was needed. No long time away from work.	Superficial wound, bruise	743	39.2
9≤ ISS ≤ 24	Medical intervention is needed. After long time there is a return to work.	Cut, fracture, trauma and internal injury	539	28.4
≥ 25	Mostly workforce loss or life loss.	Electrocution, multi-fracture, death	173	9.1

The ISS can be applied to persons who have sustained injury to more than one area of the body as well as to those with isolated injuries. Each injury is assigned an Abbreviated Injury Score (AIS) (may vary from 1-minimum severity to 6-maximum severity, almost always fatal) and is allocated to one of six body regions (head, face, abdomen, extremities, external). Only the highest AIS in each body region is used. An individual's ISS is determined by rating each injury with the AIS, then adding together the squares of the highest AIS rating for each of the three most severely injured body areas [40]. The ISS takes values from 0 to 75; it can generate only 44 values [42-44]. The ISS gives equal importance to injuries with the same AIS severity occurring in different body regions [45]. Researchers used ISS for different areas such as medicine, traffic safety [46-51]. Different categorizations were used for different cases and application areas in these studies. It is noteworthy that categorization was done in a systematic way but there was no standardization. With this in mind, the ISS calculated for the cases in this study was classified within a logical framework. ISS for each victim was estimated by using the type of injury and injured body part information found in the SSI workplace accident report forms. ISS for each casualty, ranging from 0 to 75, was calculated and then classified into 4 groups. Classification was done based on type of injury and time away from work. The framework used for ISS classifications along with the counts and percentages of analyzed cases are displayed in Table 5.

ISS, was used as a dependent quantitative variable in addition to the 15 variables grouped under four categories that are presented in Tables 1-4. ISS variable was converted into binary category at the modelling stage in the logistic regression analysis. Converting process is explained in detail in the Logistic Regression Analysis section of this paper.

4.2. Univariate Analysis

In vast majority of the construction safety literature that utilizes descriptive statistics analysis, research findings are based on univariate analysis [52-55]. Similarly, univariate analysis was adopted in this research, for two purposes: data screening and classification, and to understand what we have and choose the right variables for bivariate data analysis, which is the next step in the methodology.

As discussed in the previous section of this paper, the database created for this study consists of 15 research variables. The frequency distributions of each variable are presented in Tables 1-4. Each table and corresponding variables were analyzed and interpreted as part of univariate analysis. Results of this analysis and key findings are presented in Results and Discussion section.

4.3. Cross Tabulation Analysis

After applying univariate analysis to investigate whether there is a meaningful relationship between variable pairs, cross tabulation analysis is carried out. Cross tabulation is widely utilized to study the relationship between categorical variables [55]. Researchers used cross tabulation analysis for different purposes, such as for determination of the relationship between service quality and customer satisfaction, comparing three widely used methods for assessing alcohol consumption, tracing of the pattern of general and specific aspects of marital satisfaction over the family life cycle and determining attitudes of students towards using credit cards [56-60].

Cross tabulation analysis produces a contingency table displaying relationship, in the form of joint frequencies of two or more variables. The rows indicate one variable while the columns indicate the other [61]. The Pearson Chi-square test is one of the tests that can be used to interpret this relationship successfully [62]. Pearson chi-square compares the observed counts with those that would be expected if there were no association between two variables [63]. After calculating the Pearson Chi-squared value, the p-value based on that value (which expresses the importance of the Chi-square value) must be calculated separately. The p-value is the probability value that is used for hypothesis testing by the Pearson chi-square test. After finding the p-value, one can decide whether the result is statistically significant or not. Most common practice for significance level is 0.05, in other words the confidence interval is 95%. Therefore, a p-value less than 0.05 is accepted as significant and allows researcher to reject the null hypothesis of no association, and conclude that there is an association between variables [62].

After the null hypothesis is rejected to determine the strength of this relationship, researchers have to calculate Phi or Cramer's V values. Phi values from 0 to 0.1 show a weak relationship; 0.1 to 0.3 indicate a moderate relationship; and values between 0.3 and 1.0 suggest a strong relationship [64].

Following the conclusion of univariate frequency analysis in this study, Injury severity score (ISS) was selected as the dependent variable, while the others were assumed as independent variables. Cross tabulation analysis was performed between dependent variable ISS and each independent variable. Pearson chi-square, "p" and Phi or Cramer's V values were taken into account to statistically test the relationships of variables.

Even though, cross tabulation analysis investigates whether a significant relationship exists between the dependent and independent variables and its influence on the statistically dependent variable, the findings of the cross tabulation do not reveal information about the interactions between the independent variables. Therefore, binary logistic regression analysis was performed to check the combined effect of the statistically significant independent variables and to determine their relationships with the dependent variable.

4.4. Logistic Regression Analysis

Logistic regression (LR) is a statistical analysis tool for modeling the relationship between a response variable and a set of explanatory variables when the response variable is categorical [26]. The main goal of the LR analysis is to find the best fitting and most parsimonious, yet reasonable, model to describe the relationship between a dependent (response) and a set of independent (predictor or explanatory) variables [65]. Logistic regression does not have the requirements for the independent variables to be normally distributed, linearly related or of equal variance within each group [67]. In logistic regression, instead of predicting the value of Y (dependent variable) from predictor variable X1…n , the probability of Y occurring is predicted with given known values of $X_{1\ldots n}$ [61].

The significance of LR lies in the logistic transformation. In order to predict the dependent variable probability and, perform logistic transformation, the probability function can be written as Equation 1, where p is the probability of being in one group (occurrence of an event) and 1-p is the probability of being in the other group (non-occurrence of an event) [67].

$$\frac{p}{1-p} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon \qquad (1)$$

The problem with Equation 1 is that the right side of the equation can get any value between - ∞ to + ∞ while, the left side of the equation cannot be negative. The logit transformation equation must be used to overcome this problem, and it is formulated as the following equation where the natural log of the probability of being in one group divided by the probability of being in the other group.

$$logit\ (p) = \ln\left(\frac{p}{1-p}\right) \tag{2}$$

After logit transformation is applied the equation becomes;

$$\ln\left(\frac{p}{1-p}\right) = \ln(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon) \tag{3}$$

In solving the Equation 3, the LR equation from which the probability of Y is predicted becomes as the following equation where, P(Y) = probability of Y occurring; e is the base of natural logarithm and β_0 represents exposure variable or constant, $\beta_{1\ldots n}$ are the coefficients, and $X_{1\ldots n}$ are the independent (predictor) variables.

$$P(Y) = \frac{1}{1+e^{-(\beta_0+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_n X_n+\varepsilon)}} \tag{4}$$

In LR, to distinguish the difference between the occurrence and non-occurrence of an event, the dependent variable is coded in a certain way. Assigning a value of 1 (Y=1) to event occurrence and 0 (Y=0) to no occurrence is the simplest way to code the dependent variable [26]. In this research, injury severity score (ISS) was selected as a binary dependent variable. A value of one (1) was assigned to the variable, when ISS is ≥ 9; otherwise, it was assigned zero (0) indicating an ISS < 9. Independent variables identified as statistically significant in cross tabulation analysis were treated as multilevel independent variables.

In order to measure the performance of model's validation of the LR models is necessary.

Poorly fitting results that inaccurately predict the future outcomes may be caused doesn't apply validation to the model [68]. In this study, the model subsets were selected based on a 70/30 ratio. Bernoulli distribution was used to facilitate a random selection on injury cases. Hence Bernoulli distribution takes values of 0 and 1, 1 value was assigned randomly to 70% of the cases, which were used to develop the model and the remaining 30% was used to validate the data [69].

Initially, prediction power of the naïve model, which has only one constant (β_0) without any predictor variables, was estimated. Then, independent variables were added to the naïve model to improve its predictive power, where the improved model is called the development model in this study. A stepwise backward enter method was used during the development of the model. After incorporating the independent variables into the model, the variables that did not have significant positive effect on the predictive power of the development model were left out (assigned task, construction trade), while the significant ones remained in the model (work experience in project, type of injury, unsafe condition, and unsafe act). It should be noted that validation of 30% data set was performed for both naïve and development models.

The model creation is a challenge, to choose the best predictive model various numbers of tests should be applied. The first thing one has to do is to make sure that it meets the guidelines for "goodness-of-fit". This goodness-of fit is done by a parameter that checks the fit of the model. The log-likelihood needs to be calculated In order to do so. The log-likelihood is based on summing the probabilities associated with the predicted and actual outcomes [26]. Wald's test, Hosmer and Lemeshow's R_L and Exp (β) can be listed as other tests that need to be conducted. Wald test is used to determine whether an independent variable is a significant predictor of the outcome or not. Hosmer and Lemeshow's R_L is a test which represents the measure of how much the goodness of fit improves as a result of the inclusion of predictor variables in each step. Exp(β) is the exponential value of the β coefficients and its value represents the odds ratio. Therefore, Exp(β) represents the odds ratio of that predictor variable and how it affects the outcome [61]. Similar to the past research discussed above, this study calculated log-likelihood values, ran Wald and Hoshmer Lemeshow chi-squared tests to determine the predictive power of the development model. Interpretation of the results can be found in Results and Discussion sections.

5. RESULTS

5.1. Univariate Analysis Results

The results of univariate analysis are presented in Table 1-4, where the "count" values of every variable are displayed. This value indicates how many cases are recorded in each category, or in other words their frequency. "Percentage" value in the table represents the percentage distribution of the data categorized under the same variable type. The following inferences were drawn by analyzing the frequency distribution of each variable:

- Contrary to expectations, the number of accidents in winter turned out to be higher than in summer. This finding is quite surprising because the production output in construction industry, which is negatively affected by weather conditions in winter, usually accelerates during summer months. It was expected to see more injuries in summer due to increase in the number of workers on site. This could indicate that there are fewer risks in construction sites, while working under fair weather conditions.
- It was seen that injuries occurred on Monday the most (17.2%). Many previous research studies have proved that this finding is common, which is also named as "Monday Effect" [67-70].
- It was observed that the injuries tend to occur more between the hours of 10:00-12.00 (24.5%) and 14.00-16:00 (21.4%). It is assumed that this could happen due to acceleration of work during those time periods.
- When "type of injury" variable was examined, "falls" and "struck by falling object" were the two most noticeable data with 36.4% and 30.0% frequencies respectively. These findings show similarities with the past studies. Many researches pointed out that fall accidents are not only the highest frequent in the construction but also their results are more severe [71-76]. Similarly, Arndt et al., found that struck by falling objects is one of the common causes of fatal injuries in the construction industry [77].
- There wasn't a clear cut choice among the results of "nature of injury" variable category. Superficial wound, bruise, fracture/crack and cuts were the most encountered injuries. However, these results did not lead to a specific finding.
- Findings indicate that workers who are assigned in formwork and material handling (such as lifting and carrying materials and products) had the most number of injuries with 25.3%

and 17.5% respectively. Thus, these two construction tasks pose more risks to workers safety according to this study.

- It is observed that the worker's age group is predominantly younger. The average age of victims in the database is 33 and the workers between 19-24 years old (20.3%) had more injuries than the other age groups. It was found that the frequency of accidents decreased in proportion to the increasing age.
- Results of the study showed that work experience reduces injury severity. This somewhat anticipated finding solidifies that the increasing work experience has a positive effect on safety awareness. As the worker gains experience, they become experts in their jobs and become more aware of their surroundings with safety in mind. Safety and health training, if provided, also is a tool for reducing injury severity. Past research on similar topic also could support this finding [11].
- Educational background variable revealed an interesting finding that in more than half of the cases analyzed, workers had no further education beyond elementary school. This finding is directly associated with the fact that majority of construction force in Turkey has elementary school diploma or less.
- Similar to educational background finding, most of the cases (28.1%) involved unskilled worker because most of the construction workforce consists of unskilled laborers. However, a noteworthy amount (25.5%) of the cases involved form workers. This finding naturally matches with the findings of "assigned task" category, which is somewhat close to the description of "construction trade" variable.
- When unsafe act and condition variables were analyzed, three data categories stood out: no personal protective equipment (PPE) usage, insufficient/lack of written work procedures and unsafe work practices. Other most encountered unsafe acts and conditions were poor housekeeping, no collective protection systems, no PPE provided, and position inappropriate for the task.

5.2. Cross Tabulation Results

To investigate the relationship between ISS and other independent variables cross tabulation analysis was performed as the second part of the descriptive analysis. Only statistically significant results of cross tabulation analysis were presented in Table 6. According to the analysis, 6 of the 14 independent variables (work experience in project, assigned task, construction trade, type of injury, unsafe act, unsafe condition) were found statistically significant with $p<0.05$, as shown in Table 6.

Table 6 also displays Cramer's V values, which are used for determining the strength of relationship between variables. Findings show that work experience in project (Crv: 0.142), assigned task (Crv: 0.153), construction trade, (Crv: 0.139) and unsafe act (Crv: 0.245) have moderate relationship strength, while type of injury (Crv: 0.347) and unsafe condition (Crv: 0.330) has a strong relationship with the dependent variable (Table 6).

Table 6 - Contingency Table - ISS vs. Nominal Variables

Independent Variables	Pearson's X^2(df), p		Phi Cramer's V	
Work experience in project	$X^2(6)=27,676$	p=0.000	crv(6)=0.1420	p=0.000
Assigned task	$X^2(12)=31,112$	p=0.001	crv(12)=0.153	p=0.001
Construction trade	$X^2(7)=26,705$	p=0.000	crv(7)=0.139	p=0.000
Type of injury	$X^2(8)=166,001$	p=0.000	crv(8)=0.347	p=0.000
Unsafe Condition	$X^2(4)=149,508$	p=0.000	crv(4)=0.330	p=0.000
Unsafe Act	$X^2(4)=82,666$	p=0.000	crv(4)=0.245	p=0.000

5.3. Logistic Regression Analysis Results

Naïve and development models were generated, validated and tested for prediction power in the logistic regression analysis part of the study. Injury Severity Score (ISS) was used as the dependent variable for both models. While naïve model used six variables that were identified in cross tabulation analysis part, development model left out two variables (assigned task, construction trade) that did not have significant positive effect on the predictive power of the model. The following results were obtained from the analysis:

The naïve model shows the general percentages of the classified data without any predictor variables and uses the 70% data set, which is assumed to have ISS lower than 9. The prediction power of the naïve model was found to be 61.5%. Whereas the prediction power of the 30% data set, which is used to validate data and has ISS equal to or greater than 9, was found to be 65.4%. The development model using the 70% data set, this time with the inclusion of the predictor variables, yielded a prediction accuracy of 71.4%. The validation of the development model was performed on the 30% data set, which produced 73.1% accuracy. The outcome of logistic regression modeling is presented in Table 7.

Table 7 - Logistic Regression Modeling Results

		Naïve Model						Development Model					
		Model Development Set			Validation Set			Model Development Set			Validation Set		
		ISS			ISS			ISS			ISS		
	Observed	<9	≥9	%	<9	≥9	%	<9	≥9	%	<9	≥9	%
ISS	<9	590	0	100.0	270	0	100.0	508	82	86.1	232	38	85.9
	≥9	370	0	0.00	143	0	0.00	193	177	47.8	73	70	49.0
	Overall %			61.5			65.4			71,4			73.1

The final developed model's log likelihood value (1099.112) was found to be smaller than that of the Naïve model (1279.991), which means that the development model is more accurate in predicting the injury severity score (ISS<9 vs. ISS>9) than the Naïve model.

Finally, based on a p value of .817 from the Hoshmer Lemeshow chi-squared test, which is greater than 0.05, the development model shows a good fit with the data.

In the light of these results, it can be stated that the four significant factors (work experience in project, type of injury, unsafe act and unsafe condition) are valid predictor variables that can be used in estimating whether an accident is going to result in a high injury severity or not. The negative signs of the regression (beta) coefficients listed indicate that work experience in project has decreasing effects on the probability of a workplace injury resulting with high ISS. In other words, the more experience the worker has, the less likely he/she will be exposed to a severe workplace related injury. Faulty tool/equipment/machinery and not using Personal Protective Equipment (PPE) categories from unsafe condition and unsafe act variables respectively had the most significant effect on the ISS. Types of injury such as falls, vehicle accidents, heat exposure and electric shock had increasing effects on ISS.

6. SUMMARY AND CONCLUSION

The study presented in this paper intended to use logistic regression analysis to develop a statistical model in identifying factors that contribute to severity of construction injuries and predicting their severity scores. An extensive archival study was initially performed to establish a database to realize this study. Research methodology included using statistical analysis methods such as frequency analysis (univariate) and cross tabulation (bivariate) analyses before undertaking logistic regression.

Although, univariate and bivariate analyses were primarily used as tools to determine statistically significant variables for logistic regression, the following conclusions can be drawn from the findings:

- Cold weather conditions could have a negative impact on the frequency of injuries.
- “Monday effect” theory was once again confirmed by this study.
- “Falls” and “struck by” were identified as dominant injury types in Turkish construction industry.
- As the age of worker increases, it was observed that frequency of work related injuries decreases. This could be an indicator of work experience positively affecting safety awareness.
- Formwork workers tend to have more injuries than other worker trades.
- Unsafe act and condition variables revealed that safety management or lack thereof is a major problem in Turkish construction sites. Particularly insufficient or unsafe work procedures and not using PPE were encountered the most in accident reports.

  The outcome of the logistic regression analysis showed that four of the independent variables (work experience, accident type, unsafe condition and unsafe act) have statistically significant influence on workplace injury severity. To elaborate further on the regression results;

- Workers who have more work experience than others are less likely to be exposed to a severe workplace related injury.

- Faulty tool/equipment/machinery and not using Personal Protective Equipment (PPE) categories from unsafe condition and unsafe act variables respectively had the most significant effect on the severity of construction injuries. In other words, injuries caused by those unsafe acts and conditions are more likely to end up severe.
- Types of injury such as falls, vehicle accidents, heat exposure and electric shock have higher probability of causing severe injuries.

7. DISCUSSION AND RECOMMENDATIONS

Injury severity score (ISS), which had not been applied in the construction safety related research studies before, was identified as the dependent variable in the logistic regression analysis part of the study. Using ISS values as a dependent variable and including scoring in the model has been one of the features that made this study unique. In addition, the nonlinear nature of the ISS allowed the use of this scoring system as a dependent variable in the logistic regression analysis, which has a similar nonlinear structure. Using ISS makes it possible to compare and rank injuries by their severity and could be used in future studies that involve estimating the costs of injuries and risk analysis.

Research information on modeling the accident outcomes is scarce in safety and health literature. In fact, this was one of the main driving forces of this study that led to selection of binary logistics regression analysis, which is usually used in research areas other than construction safety. Logistic regression analysis developed a model that consists of four multilevel independent variables (work experience in project, type of injury, unsafe condition, and unsafe act), all of which have significant relationships with the binary dependent variable (injury severity score). The improvement of accuracy from the naïve to the development model (9.9%), and further improvement of accuracy on the validation model (7.7%) are indicative of the successful validation of the logistic regression model presented.

The model developed in the study allows to predict the injury severity (whether the ISS is greater or less than 9) by using the information coming from independent variables included in the model and to approach proactively. It would be possible to at least reduce the severity of potential injuries by focusing on the factors that meaningfully contribute to the outcome of injury; such as, increasing training efforts for workers with less work experience, eliminating fall and struck by falling object hazards, providing PPE and encouraging their usage, enforcing written safety procedures.

This study was realized by investigating more than two thousand construction accident case reports that occurred in three big cities (İstanbul, Ankara and İzmir) of Turkey over a span of three years with the special permission of Social Security Institution. It is expected that the database created for this study could be useful in other statistical studies involving Turkey, where accurate statistical injury data is still difficult to obtain. Furthermore, a similar study could be carried out in different locations or with different sectors and comparisons could be made with the findings of this study.

References

[1] Aksorn, T., Hadikusumo, B.H.W. Critical success factors influencing safety program performance in Thai construction projects. Safety Science, 46, 709-727, 2008.

[2] Zeng, S.X., Vivian, W.Y., Tam, C.M. Towards occupational health and safety systems in the construction industry of China. Safety Science, 46, 1155-1168, 2008.

[3] Gillen, M., Baltz, D., Gassel, M., Kirsch, L., Vaccaro, D. Perceived safety climate, job demands, and coworker support among union and nonunion injured construction workers. Journal of Safety Research 33, 33-51, 2002.

[4] Fass, S., Yousef, R., Liginlal, D., Vyas, P., Understanding causes of fall and struck-by incidents: What differentiates construction safety in the Arabian Gulf region?, Applied Ergonomics, 58, 515-526, 2017.

[5] Pérez-Alonso, J., Carreno-Ortega, A., Vázquez-Cabrera, F.J., Callejón-Ferre, A.J. Accidents in the greenhouse-construction industry of SE Spain. Applied Ergonomics, 43, 69-80, 2012.

[6] Unsar, S., Sut, N. General assessment of the occupational accidents that occurred in Turkey between the years 2000 and 2005. Safety Science, 47, 614-619, 2009.

[7] Haslam, R.A., Hide, S.A., Gibb, A.G.F., Gyi, D.E., Pavitt, T., Atkinson, S., Duff, A.R. Contributing factors in construction accidents. Applied Ergonomics, 36(4), 401-415, 2005.

[8] Çolak, B., Etiler, N., Biçer, U. Fatal Occupational Injuries in the Construction Sector in Kocaeli, Turkey, 1990-2001. Industrial Health, 24, 424-430, 2004.

[9] Spangenberg, S., Baarts, C., Dyreborg, J., Jensen, L., Kines, P., Mikkelsen, K.L. Factors contributing to the differences in work related injury rates between Danish and Swedish construction workers. Safety Science 41, 517-530, 2003.

[10] Tözer, K.D., Çelik, T., Gürcanlı, G.E., Classification of Construction Accidents in Northern Cyprus. Teknik Dergi, DOI: 10.18400/tekderg.325546, 29 (2), 8295-8316, 2018.

[11] Bilir, S., Gürcanli, G.E., A Method For Determination of Accident Probability in Construction Industry. Teknik Dergi, DOI: 10.18400/tekderg.363613, 29 (4), 8537-8561, 2018.

[12] Baradan, S., Akboga, O., Cetinkaya, U., Usmen, M., Ege Bölgesindeki İnşaat İş Kazalarının Sıklık ve Çapraz Tablolama Analizleri. Teknik Dergi, 27 (1), 7345-7370, 2016.

[13] Al-Ghamdi, A.S. Using Logistic Regression to Estimate the Influence of Accident Factors on Accident Severity. Accident Analysis and Prevention, 34, 729-741, 2002.

[14] Mercier, C.R., Shelley, II M.C., Rimkus, J.B., Mercier, J.M. Age and Gender as Predictors of Injury Severity in Head-on Highway Vehicular Collisions. Transportation Research Record: Journal of the Transportation Research Board, 1581, DOI: 10.3141/1581-05, 1997.

[15] Kim, K., Li, L., Modeling Fault among Bicyclists and Drivers Involved in Collisions in Hawaii, 1986–1991. Transportation Research Record: Journal of the Transportation Research Board, 1538, 75-80, 1996.

[16] Steyeberg, E.W., Harrell, Jr F.E., Borsboom, G.J., Eijkemans, M.J., Vergouwe, Y., Habbema, J.D. Internal Validation of Predictive Models: Efficiency of Some Procedures for Logistic Regression Analysis. Journal of Clinical Epidemiology, 54: 774-781, 2001.

[17] Li, Y., Bai, Y. Development of Crash-Severity-Index Models for the Measurement of Work Zone Risk Levels. Accident Analysis and Prevention, 40, 1724-1731, 2008.

[18] Fang, D., Chen, Y., Wong, L. Safety Climate in Construction Industry: A Case Study in Hong Kong. Journal of Construction Engineering and Management, 132, 573-584 2006.

[19] Cakan, H., Kazan, E., Usmen, M. Investigation of Factors Contributing to Fatal and Nonfatal Roofer Fall Accidents. International Journal of Construction Education and Research, 10, 300-317, 2014.

[20] Irizarry, J., Abraham, D.M. Assessment of Risk Perception of Ironworkers. Journal of Construction Research, 7(1/2), 111-132, 2006.

[21] Baradan, S., Usmen, M. Comparative Injury and Fatality Risk Analysis of Building Trades. Journal of Construction Engineering and Management, 132 (5), 533-539, 2006.

[22] Social Security Institution (SSI) Official Statistics http://www.sgk.gov.tr/wps/portal/sgk/tr/kurumsal/istatistik/sgk_istatistik_yilliklari Accessed February 07, 2018.

[23] Hürriyet, Daily News, 10 workers killed in construction site accident in Istanbul, http://www.hurriyetdailynews.com/10-workers-killed-in-construction-site-accident-in-istanbul-.aspx?pageID=238&nID=71381&NewsCatID=341 Accessed February 02, 2018.

[24] The Telegraph, At least 11 die in fire at construction site in Turkey, http://www.telegraph.co.uk/news/worldnews/europe/turkey/9137516/At-least-11-die-in-fire-at-construction-site-in-Turkey.html Accessed February 01, 2018.

[25] Act No 6331, Occupational Health and Safety Law, Date of Enactment: 20/06/2012.

[26] Tabachnick, B.G., Fidell, L.S. Using Multivariate Statistics. 6th Ed., Pearson Education. Inc., Boston, MA. 2013.

[27] Kass, G.V. An exploratory technique for investigating large quantities of categorical data. Applied Statistics, 29:2, 119–127, 1980.

[28] Suruda, A., Fosbroke, D., Braddee, R. Fatal work-related falls from roofs. Journal of Safety Researc, 26(1), 1-8, 1995.

[29] Stern, F.B., Ruder, A.M., Chen, G. Proportionate mortality among unionized roofers and waterproofers. American Journal of Industrial Medicine, 37(5), 478-492, 2000.

[30] Huang, X., Hinze, J. Analysis of Construction Worker Fall Accidents. Journal of Construction Engineering and Management, 123(3), 262-271, 2003.

[31] Beavers, J.E., Moore, J.R., Schriver, W.R. Steel Erection Fatalities in the Construction Industry. Journal of Construction Engineering and Management, 227-234, 2009.

[32] Im, H.J., Kwon, Y.J., Kim, S.G., Kim, Y.K., Ju, Y.S., Lee, H.P. The characteristics of fatal occupational injuries in Korea's Construction Industry. Safety Science, 47, 1159-1162, 2009.

[33] Gürcanlı, E., Müngen, U. Analysis of construction accidents in Turkey and responsible parties. Industrial Health, 51, 581-598, 2013.

[34] ICD, International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10) Version for 2010, Chapter XX External causes of morbidity and mortality (V01-Y98), W00-X59 Other external causes of accidental injury, http://apps.who.int/classifications/icd10/browse/2010/en#/XX Accessed May 03, 2018.

[35] ILO, 2014, "Theory of Accident Causes", http://www.iloencyclopaedia.org/part-viii-12633/accident-prevention/92-56-accident-prevention/theory-of-accident-causes Accessed April 21, 2018.

[36] Hill, D.C. Construction Safety Management and Engineering. American Society of Safety Engineers, Des Plaines, Illinois, 707 p. 2004.

[37] DMI, Disaster Management Institute, "The Domino Theory", http://www.hrdp-idrm.in/e5783/e17327/e24075/e27357/ Accessed January 24, 2017.

[38] Linn, S. The Injury Severity Score-Importance and Use. Ann Epidemiol., 5(6), 440-446, 1995.

[39] Baker, S.P. Evaluation of Medical Care of the Injured. Journal of Trauma, 11, 892-894, 1971.

[40] FTR, Florida Trauma Registry, A Report on The Collection Of Trauma Patient-Level Data From Trauma Centers And Other Acute Care Facilities For 2002. 2002 Annual Report, Division of Emergency Medical Operations Florida Department of Health July, 2003.

[41] Baker, S.P., O'Neill, B. The Injury Severity Score: An Update. The Journal of Trauma, 16(11), 882-885, 1976.

[42] Krischer, J.P. Indexes of Severity: Underlying concepts. Health Serv. Res., 11, 143-157, 1976.

[43] Krischer, J.P. Indexes of Severity: "Conceptual Development. Health Serv. Res., 14, 56-57, 1979.

[44] Sacco, W.J., Jameson, J.W., Copes, W.S., Lawnick, M.M., Keast, S.L., Champion, H.R. Progress toward a New Injury Severity Characterization: Severity Profiles. Comput. Biol. Med., 18, 419-428, 1988.

[45] Copes, W.S., Champion, H.R., Sacco, W.J., et al., Progress in Characterizing Anatomic Injury. Journal of Trauma, 30, 1200-1207, 1990.

[46] Rainer, T.H., Chan, S.Y., Suen, D.T.K., Lam, W., Cocks, R.A. Severe Trauma Presenting to the Resuscitation Room of a Hong Kong Emergency Department. Hong Kong Journal of Emergency Medicine, 7(3), 129-135, 2000.

[47] Ballesteros, M.F., Dischinger, P.C., Langenberg, P. Pedestrian Injuries and Vehicle Type in Maryland, 1995-1999. Accident Analysis and Prevention, 36, 73-81, 2004.

[48] Moore, S.J. The Effects of Daily Snowfall Accumulation on the Pattern and Severity of Traumatic Injuries from a U.S. Ski Resort. Trauma/Critical care Services, Vail Valley Medical Center, Colorado, 2011.

[49] Moylan, J.A., Detmer, D.E., Rose, J., et al:, Evaluation of the Quality of Hospital Care for Major Trauma. Journal of Trauma, 16, 517-523, 1976.

[50] Gillen, M., Faucett, J.A., Beaumont, J.J., McLoughlin, E. Injury Severity Associated with Nonfatal Construction Falls. American journal of Industrial Medicine, 32, 647-655, 1997.

[51] Glancy, K.E., Glancy, C.J., Lucke, J.F., et Al, A Study of Recovery in Trauma Patients. F., Trauma, 33, 602-609, 1992.

[52] Richardson, D., Loomis, D., Bena, J., Bailer, A.J. Fatal occupational injury rates in southern and non-southern states, by race and Hispanic ethnicity. Am. J. Public Health 94, 1756–1761, 2004.

[53] Ringen, K., Englund, A., Welch, L., Weeks, J.L., Seegal, J.L. Why construction is different. Occupational Medicine: State of the Art Reviews, 10(2), 255– 259, 1995.

[54] Ringen, K., Seegal, J.L., Englund, A. Safety and health in the construction industry. American Journal of Public Health, 16, 165–188, 1995.

[55] Babbie, E.R. The Basics of Social Research. 5th Ed., Wadsworth Publishing, California, 2011.

[56] Greenland, S. Modeling and variable selection in epidemiologic analysis. American Journal of Public Health, 79(3), 340- 349, 1989.

[57] Rehm, J., Greenfield, T.K., Walsh, G., Xie, X., Robson, L., Single, E. Assessment methods for alcohol consumption, prevalence of high risk drinking and harm: a sensitivity analysis. International Journal of Epidemiology, 28(2), 219-224, 1999.

[58] Arias, J.C., Miller, R. Market analysis of student' attitudes about credit cards. Business Intelligence Journal, 3(1), 23-36, 2010.

[59] Rollins, B.C., Feldman, H. Marital Satisfaction over the Family Life Cycle. Journal of Marriage and Family, 32(1), 20-28, 1970.

[60] Sureshchandar, G.S., Rajendran, C., Anantharaman, R. The Relationship between Service Quality and Customer Satisfaction–A Factor Specific Approach. Journal of Services Marketing, 16(4), 363-379, 2002.

[61] Fields, A. Discovering Statistics Using SPSS. 2nd Ed., Sage Publications, Thousand Oaks, CA, 2005.

[62] Sims, R.L. Bivariate Data Analysis: A Practical Guide. Nova Science Pub., Inc. 1999.

[63] Elliott, A.C., Woodward, W.A. Statistical Analysis Quick Reference Guidebook with SPSS Examples. 1st Ed., SAGE Publications, Inc., California. 2007.

[64] Healy, J. Statistics: a tool for social research. Wadsworth Publishing, 9th edition. 2011.

[65] Hosmer, W.D., Lemeshow, S. Applied logistic regression. John Wiley & Sons, Hoboken, NJ. 2000.

[66] Kleinbaum, D.G. Logistic regression: A self-learning text. Springer-Verlag; 1st edition, New York, NY. 1994.

[67] Kazan, E.E. Analysis of Fatal and Nonfatal Accidents Involving Earthmoving Equipment Operators and On-Foot Workers. PhD Thesis, Wayne State Universty, 173p. 2013.

[68] Arboretti Giancristofaro, R., Salmaso, L. Model Performance Analysis and Model Validation in Logistic Regression. Statistica LXIII, 375 – 396, 2003.

[69] Azen, R., Walker, C.M. Categorical Data Analysis for the Behavioral and Social Sciences. 1st Ed., Routledge Academic, New York. 2010.

[70] Campolieti, M., Hyatt, D.E. Further Evidence on the "Monday Effect" in Workers' Compensation. Industrial and Labor Relations Review", 59(3), 438-450, 2006.

[71] Chi, C.F., Chang, T.C., Ting, H.I. Accident patterns and prevention measures for fatal occupational falls in the construction industry. Applied Ergonomics, 36 (4); 391-400, 2005.

[72] Kemmlert, K., Lundholm, L. Slips, trips and falls in different work groups—with reference to age and from a preventive perspective. Applied Ergonomics, 32(2), 149-153, 2001.

[73] Ore, T., Stout, N. Traumatic occupational fatalities in the US and Australian construction industries. Am J Ind Med 30, 202–206, 1996.

[74] Hinze, J., Pedersen, C., Fredley, J. Identifying root causes of construction injuries. Journal of Construction Engineering and Management, 124(1), 67-71, 1998.

[75] Jackson, S.A., Loomis, D. Fatal Occupational Injuries in the North Carolina Construction Industry, 1978–1994. Applied Occup Environ Hyg 17, 27–33, 2002.

[76] Fabrega, V., Stakey, S. Fatal Occupational Injuries among Hispanic Construction Workers of Texas, 1997 to 1999. Human and Ecological Risk Assessment, 7, 1869–1883, 2001.

[77] Arndt, V., Rothenbacher, D., Daniel, U., Zschenderlein, B., Schuberth, S., Brenner, H. All-cause and cause specific mortality in a cohort of 20,000 construction workers; results from a 10 year follow up. Occup. Environ. Med. 61(5), 419-425, 2004.

TEKNİK DERGİ MANUSCRIPT DRAFTING RULES

1. The whole manuscript (text, charts, equations, drawings etc.) should be arranged in Word and submitted in ready to print format. The article should be typed on A4 (210 x 297 mm) size paper using 10 pt (main title 15 pt) Times New Roman font, single spacing. Margins should be 40 mm on the left and right sides and 52.5 mm at the top and bottom of the page.
2. Including drawings and tables, articles should not exceed 25 pages, technical notes 6 pages.
3. Your contributed manuscript must be sent over the DergiPark system. (http://dergipark.gov.tr/tekderg)
4. The text must be written in a clear and understandable language, conform to the grammar rules. Third singular person and passive tense must be used, and no inverted sentences should be contained.
5. Title must be short (10 words maximum) and clear, and reflect the content of the paper.
6. Sections should be arranged as: (i) abstract and keywords, (ii) title, abstract and keywords in the other language, (iii) main text, (iv) symbols, (v) acknowledgements (if required) and (vi) references.
7. Both abstracts should briefly describe the object, scope, method and conclusions of the work and should not exceed 100 words. If necessary, abstracts may be re-written without consulting the author. At least three keywords must be given. Titles, abstracts and keywords must be fitted in the first page leaving ten line space at the bottom of the first page and the main text must start in the second page.
8. Section and sub-section titles must be numbered complying with the standard TS1212.
9. Symbols must conform to the international rules; each symbol must be defined where it appears first, additionally, a list of symbols must be given in alphabetic order (first Latin, then Greek alphabets) at the end of the text (before References).
10. Equations must be numbered and these numbers must be shown in brackets at the end of the line.
11. Tables, drawings and photographs must be placed inside the text, each one should have a number and title and titles should be written above the tables and below the drawings and photographs.
12. Only SI units must be used in the manuscripts.
13. Quotes must be given in inverted commas and the source must be indicated with a reference number.
14. Acknowledgement must be short and mention the people/ institutions contributed or assisted the study.
15. References must be numbered (in brackets) in the text referring to the reference list arranged in the order of appearance in the text. References must include the following information:

    <u>If the reference is an article:</u> Author's surname, his/her initials, other authors, full title of the article, name of the journal, volume, issue, starting and ending pages, year of publication.

        Example : Naghdi, P. M., Kalnins, A., On Vibrations of Elastic Spherical Shells. J. Appl. Mech., 29, 65-72, 1962.

    <u>If the reference is a book:</u> Author's surname, his/her initials, other authors, title of the book, volume number, editor if available, place of publication, year of publication.

        Example : Kraus. H., Thin Elastic Shells, New York. Wiley, 1967.

    <u>If the reference is a conference paper:</u> Author's surname, his/her initials, other authors, title of the paper, title of the conference, location and year.

    <u>If the source is a thesis:</u> Author's surname, his/her initials, thesis title, level, university, year.

    <u>If the source is a report:</u> Author's surname, his/her initials, other authors, title of the report, type, number, institution it is submitted to, publication place, year.
16. Discussions to an article published in Teknik Dergi should not exceed two pages, must briefly express the addressed points, must criticize the content, not the author and must be written in a polite language. Authors' closing remarks must also follow the above rules.
17. A separate note should accompany the manuscript. The note should include, (i) authors' names, business and home addresses and phone numbers, (ii) brief resumes of the authors and (iii) a statement "I declare in honesty that this article is the product of a genuinely original study and that a similar version of the article has not been previously published anywhere else" signed by all authors.
18. Copyright has to be transferred to UCTEA Turkish Chamber of Civil Engineers. The standard copyright form signed by the authorised author should therefore be submitted together with the manuscript.

CONTENTS

ISSN: 1300-3453