ISSN : 2148-2225

alphanumeric

The Journal of
Operations Research, Statistics,
Econometrics and Management Information Systems

Volume 4
Issue 1
2016

Peer Reviewed International Journal

http://www.alphanumericjournal.com/

ALPHANUMERIC JOURNAL

Volume 4, Issue 1, 2016

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

ISSN 2148-2225

http://www.alphanumericjournal.com/

http://dergipark.ulabim.gov.tr/alphanumeric

Alphanumeric Journal is a "Peer-Reviewed International Journal".

The journal is semiannually published

in June and December.

Alphanumeric Journal indexed by,

and accessable from,

ALPHANUMERIC JOURNAL

Volume 4, Issue 1, 2016

EDITORIAL BOARD

This page intentionally left blank

ALPHANUMERIC JOURNAL

Volume 4, Issue 1, 2016

CONTENTS

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2016.04.01.STAT.01

ASSESSING THE EFFECTS OF VARIOUS SOCIO-ECONOMIC AND HEALTH INDICATORS ON HDI COUNTRY CATEGORIES

Özlem YORULMAZ*

Assist. Prof. Dr., Faculty of Economics, Istanbul University, Istanbul

Received: 27 January 2016

Accepted: 03 March 2016

Abstract

The human development indicator (HDI) is based on three indicators: standard of living, life expectancy, and education level. Although being widely known and commonly used, the accuracy of the HDI has been criticized in the literature due to the inadequacy of its indicators. The present study uses 11 indicators to classify countries and compares the results by country groups against similar HDI ranked country groups. Furthermore, using multinomial logistic regression analysis, the effects of the 11 indicators on the country categories of HDI are investigated. The findings show that although the main cluster characteristics are similar to HDI categories, some differences exist in the classification of countries. Health indicators have a striking effect on low HDI countries relative to high HDI countries. FDI inflows and CO2 emissions per capita are significant indicators for low and middle HDI relative to high HDI countries. However women's involvement in parliament and work are not distinctive or effective indicators.

Keywords: Cluster Analysis, Gender Effect, Health Indicators, Human Development Indicator, Multidimensional Scaling, Multinomial Logistic Regression

ÇEŞİTLİ SOSYO-EKONOMİK VE SAĞLIK GÖSTERGELERİNİN İNSANİ GELİŞME ENDEKSİ ÜLKE KATEGORİLERİ ÜZERİNDEKİ ETKİLERİNİN DEĞERLENDİRİLMESİ

Özet

İnsani Gelişme Endeksi (İGE) yaşam standardı, ortalama yaşam beklentisi ve eğitim düzeyi göstergelerine dayalıdır. İGE'nin tanınırlığı ve yaygın kullanımına karşın, endeksin yetersiz faktörlerle açıklanmasından dolayı doğruluğu konusu literatürde tartışılmaktadır. Bu çalışma 11 göstergeyi dikkate alarak ülkeleri kümelemiş ve elde edilen kümeleri İGE ülke kategorileri ile karşılaştırmıştır. Ayrıca söz konusu 11 faktörün İGE ülke kategorileri üzerindeki etkisi multinomial lojistik regresyon analizi ile değerlendirilmiştir. Elde edilen bulgular ülke sınıflarının genel özelliklerinin İGE kategorileriyle benzerlik gösterdiğini ama bununla beraber sınıflamada bazı farklılıkların olduğunu göstermiştir. Sağlık göstergeleri, düşük İGE kategorisinde yer alan ülkelerde, çok yüksek İGE kategorinde yer alan ülkelere oranla oldukça etkilidir. Yüksek IGE kategorisinde yer alan ülkeler referans olarak alındığında, doğrudan yabancı yatırımlar ve kişi başına düşen CO2 emisyonu göstergelerinin, düşük ve orta İGE kategorisinde yer alan ülkelerde istatistiksel olarak anlamlı olduğu görülmüştür. Bununla beraber, kadınların parlamentodaki ve iş gücündeki katılımları etkili ya da ayırt edici bir faktör değildir.

Anahtar Kelimeler : Insani Gelişme Endeksi, Sağlık Göstergeleri, Cinsiyet Etkisi, Multinomial Lojistik Regresyon Analizi, Kümeleme Analizi, Çok Boyutlu Ölçekleme Analizi

* yorulmaz@istanbul.edu.tr (Corresponding author)

1. INTRODUCTION

With respect to countries, the concept of development allows us to distinguish between rich and poor. It is related to "...shares of resources used to provide free health and education services, equitable distribution of income among social groups, effects of production and consumption on people's environment" (Soubbotina and Sheram, 2000). Gross domestic product (GDP) is an overly simplistic measure of development and the results in an inaccurate representation of actual human processes (Barreiro, 2006). However, for a long time, per capita income has been widely used as a means of making comparisons between the development indices of countries. If we look closer at countries with similar per capita GDP, it is possible to determine differences in the availability of clean air and water, conditions of education and health care systems, the unemployment ratio, the ratio of women's participation in parliament, and flows of foreign direct investment (FDI), among other factors. The objective of development is to create an enabling environment for people to enjoy long and creative lives (UNDP, 1990).

Researchers and policymakers have long investigated economic, social, political, and environmental indicators with which to profile countries, identify development policies and to rank their state of development. Before the 1990s, economic growth was commonly used as an indicator of development. Over time, economic growth was incorporated into multidimensional measures, that included several aspects of well-being, going beyond per capita income and GDP growth. Sen (1992) introduced the capability approach, which initiated the design of the human development index (HDI). However, term HDI did not appear in the literature until the publication of Human Development Report (1990), which at this point comprised a simple unweighted average of a nation's longevity, education, and income. Over time, modifications were made to the HDI. According to the Human Development Report published in 2014, the definition of HDI is a summary measure of achievements based on key dimensions of human development: a long and healthy life, access to knowledge, and a decent standard of living (Figure 1). Consequently, the HDI has become the geometric mean of normalized indices for each of these three dimensions.

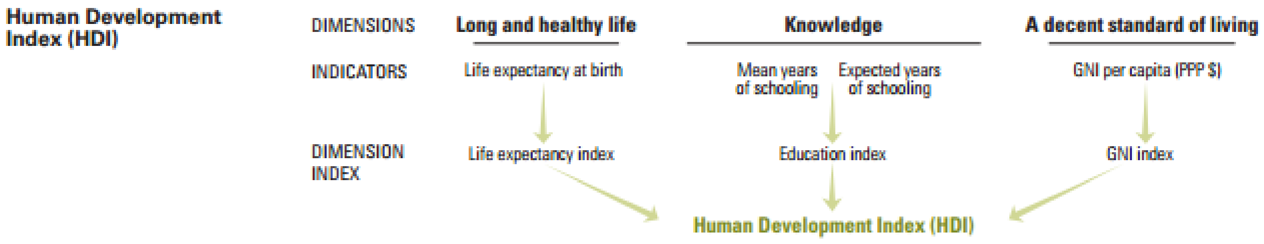


Figure 1. Sub-dimensions of HDI. Source: United Nations Report, 2014

The 2014 Human Development Report classified countries into four groups. The first group corresponds to "very high human development" with HDI values of 0.800-1.000. HDI values of 0.7-0.799 refer to "high human development" whereas values of 0.550-0.699 refer to "medium human development". The last group "low human development" has an HDI values falling below 0.55.

Despite its widespread use, the HDI has fallen under some criticism in the development literature. Most of these criticisms have focused upon the few arbitrary indicators comprising the HDI's sub-dimensions. Considering the data collection problems that some countries encounter, Srinivasan (1994) argues that "the HDI is conceptually weak and empirically unsound". Ranis et al. (2006) investigated the effects of 39 indicators and concluded that the HDI represented an incomplete measure. Wolff et al. (2010) argued that classifications based on the HDI were not reliable due to the presence of errors in the data, particularly in the health and education statistics.

The aim of this study is to investigate the effects of 11 socio-economic and health indicators on the HDI of country categories by way of multinomial logistic regression analysis. Additionally, this study examines the classification of countries in consideration of these indicators and evaluates the main characteristics of these groups by cluster analysis. The representation of HDI sub-dimension indicators is assessed based on the findings of the cluster analysis. To the best of our knowledge, this is the first study that has examined the effects of various indicators on HDI country categories. The results can be used to better inform policymakers on the shortcomings of countries ranked using the HDI and thus serve as a guide for decision making.

The remainder of the paper is organized into four sections. Section 2 reviews the relevant literature. Section 3 explains the methodology used throughout this study. Section 4 presents the data and findings, and Section 5 provides a conclusion.

2. LITERATURE REVIEW

Since the development of the HDI by the United Nations Development Programme in its 1990 report, the measure has been integrated in a wide range of studies. HDI embodies Sen's "capabilities" approach to understanding human well-being. Capabilities are instrumentalized in HDI as access to health, education, and goods (Stanton, 2007). However, as mentioned in the previous section, Ranis et al. (2006) investigated the effects of 31 indicators on human development and found HDI is an incomplete measure. In this section, we discuss several earlier studies with respect to the effects of socio-economic, demographic, and environmental variables on human development.

Lee et al. (1997) studied how HDI can be used to predict the infant and maternal mortality rate and reported HDI as a powerful predictor. Antony, Visweswarraro, and Balakrishna (1999) evaluated the representativeness of HDI for health and nutrition indicators of 174 countries. They found that dietary indicators are substantial for classification and that countries with high HDI have high rates of female education and lower rates of fertility and infant and maternal mortality. Self and Grabowski (2003) found that the contribution of health care expenditure as a percentage of GDP changes in developing and less developed countries. Ranis, Stewart and Somma (2005) found that under-five mortality rates are a good indicator of HDI. Similarly, Boutayeb and Serghihi (2006) showed that deficiencies of health (maternal and infant mortality rates) impede human development in the majority of Arab countries.

Sharma (1997) and Fukuda-Parr (2001) considered human development from a gender perspective and the importance of women on development level. Sharma (1997) argued that HDI could be improved by using sex-disaggregated data to examine the contributions various socio-economic factors.

Costa, Reybski, and Knopp (2011) found a positive relationship between per capita CO2 emissions for developing countries. Bedir and Yilmaz (2015) investigated the casual relationship between CO2 emissions and HDI for OECD countries and found a strong effect of CO2 on HDI in some of these countries.

Sharma, and Gani (2004) examined the effect of foreign direct investment (FDI) on HDI for middle- and low-income countries and found a positive effect. Later, Reiter and Steensma (2010) also found that FDI inflows affect the improvement of human development.

Saito (2003) emphasised the importance of education for human development. Njoh (2003) explored the relationship between urbanisation and the HDI, finding a positive correlation between the two for sub-Sahara Africa. Using discriminant analysis, Öztürk (2007) classified countries and predicted country categories through socioeconomic, education, and health indicators. He concluded that developed and undeveloped countries could be distinguished by their health indicators. The study also considered the proportion of seats in parliament, percentage of school enrolment and percentage of trade, concluding that these indicators could also be used to distinguish between developing and developed countries.

3. METHODOLOGIES

In this study, we employed cluster analysis to classify countries according to a range of socio-economic and health indicators. Multidimensional Scaling (MDS) was subsequently used to explore the links between the link between the indicators. In the final stage of analysis, we used multinomial logistic regression to examine the effects of the indicators on the HDI country categories.

3.1. Cluster Analysis

Cluster analysis comprises a range of methods used to identify homogeneous groups. As such, cluster analysis is used to ensure similarity within groups and, as much as possible, differences between groups. Three clustering techniques are defined in the literature: agglomerative hierarchical clustering, K-means, and density-based spatial clustering of applications with noise (DBSCAN). In our study, only hierarchical and K-means clustering were employed. In the K-means approach, the number of clusters is defined by researchers. At the beginning of the algorithm, K centroids are chosen and each observation is assigned to the closest centroid. The group of observations assigned to a cluster is called a cluster. With the assigned observations, the centroid of clusters is updated. The assignment procedure is repeated until the same centroids are obtained. The mentioned centroids are representative of each cluster's prototype. The objective functions and the centroids are outlined in Table 1.

Two approaches to hierarchical clustering are defined in the literature, namely, agglomerative and divisive. Compared to the divisive approach, the agglomerative procedure has received more attention in applications (Tan, Steinbach, & Kumar, 2005). The agglomerative approach begins with each observation as a singleton cluster and then, at each step, the two closest clusters are merged according to their proximities. The procedure continues until one cluster remains. Contrary to the agglomerative approach, divisive clustering begins with one cluster, which includes all observations. At each step, clusters are split until only singleton clusters of individual observations remain. The proximities used in agglomeration are calculated using different approaches, namely, single linkage (nearest neighbor), complete linkage (furthest neighbor), average linkage, centroid, and Ward's method (Sarstedt & Mooi, 2014).

Table 1. Objective Function and Centroid Choices in K-means Cluster Analysis

Proximity Function	Centroid	Objective Function
Manhattan (L_1)	Median	Minimize sum of L_1 distance of an object to its cluster centroid
Squared Euclidean (L_2)	Mean	Minimize sum of the squared L_2 distance of an object to its cluster centroid
Cosine	Mean	Maximize sum of the cosine similarity of an an object to its cluster centroid
Bregman divergence	Mean	Minimize sum of the Bregman divergence of an object to its cluster centroid

*Source: Tan, Steinbach, and Kumar (2005, p.501).

3.2. Multidimensional Scaling (MDS) Analysis

This technique determines the distances between objects and in order to provide a visual representation of objects in a low-dimensional space. Distances are referred to as proximities and proximity measures are described as dissimilarities, similarities, and correlations. Euclidean, Mahalanobis, quadratic Euclidean, Chebychev, Block, and Minkowski distances are defined in the literature. However, Euclidean distance has been preferred frequently in studies.

The Euclidean distance between the ith and jth points, d_{ij} is defined as

$d_{ij}^2 = \sum_{k=1}^{p} (x_{ik-} x_{jk})^2$ where p is the dimension of the observations.

The steps of the algorithm can be summarized as (Hintze, p.435):

1. Using Euclidean distances, elements of distance matrix D are calculated for observations.
2. Elements of matrix A, $A = \{-\frac{1}{2} d_{ij}^2\}$ are calculated by means of matrix D. Using matrix A, $B = \{a_{ij} - a_{i.} - a_{.j} + a_{..}\}$ is obtained where $a_{i.}$ is the average of all a_{ij} across j. The p largest eigenvalues $\lambda_1 > \lambda_2 > ... > \lambda_p$ and eigenvectors $L = (L_{(1)}, L_{(2)}, ..., L_{(p)})$ of matrix B are obtained and normalized so that $L_{(i)}^{'} L_{(i)} = \lambda_i$.
3. The rows of L refer to coordinates of the objects. $\hat{d}_{ij}$ distances from L are calculated.

To evaluate how well the dataset is represented by MDS, a goodness of fit statistic is defined for the actual values and their predicted values:

$$Stress = \sqrt{\frac{\sum\sum_{i<j}\left(d_{ij} - \hat{d}_{ij}\right)^2}{\sum\sum_{i<j}\left(d_{ij}\right)^2}}$$

The goodness of fit measure provides the closeness between original distances and predicted distances and is known as stress. Stress values < 0.05 define perfect fit.

3.3. Multinomial Logistic Regression

A multinomial logistic model determines the effects of explanatory variables on a subject chosen from a discrete set of options (Agresti, 2002). The result is a generalized binary model, the response variable having more than two categories.

Consider u possible outcomes {1,2,…,u} with multinomial probabilities $P[y = k] = p_k$, k=1,2,…,u. The probabilities can be parametrized as (Ledolter, 2013, p.132):

$$p_1 = P[y=1] = \frac{1}{1 + \sum_{h=2}^{u} \exp(\alpha_h + x\beta_h)} \quad (1)$$

$$p_k = P[y=k] = \frac{\exp(\alpha_k + x\beta_k)}{1 + \sum_{h=2}^{u} \exp(\alpha_h + x\beta_h)} \quad \text{for } k=2,\dots,u.$$

The sum of probabilities is equal to 1. The interpretation of the odds-ratio is based on following equation"

$$\log\left(\frac{p_k}{p_1}\right) = \alpha_k + x\beta_k \quad \text{for } k=2,\dots,u. \quad (2)$$

The maximum likelihood estimation approach was used to estimate parameters. Multinomial logistic regression is similar to discriminant analysis. Discriminant analysis employs a regression line to separate sample into groups whereas multinomial logistic regression analysis uses probabilities and u-1 log odds equations to determine categories. Assumptions about multivariate normality and homoscedasticity are required for discriminant analysis; however, multinomial logistic assumes neither normality nor homoscedasticity

4. DATA AND FINDINGS

This study aimed to investigate the effects of selected socioeconomic and health indicators on HDI country classifications and country categories using data sourced from the 2013 United Nations Development Report. Data from 175 countries were considered as variables for the purposes of analysis, including: Maternal Mortality Rate (V1), Adolescent Birth Rate (V2), Women Share of Seats in Parliament (V3), Female Labor Force Participation Rate (V4), Deaths Due to Tuberculosis (V5), Increase in Rate of Physicians (V6), Public Health Expenditure (% of GDP) (V7), CO2 Emissions per Capita (V8), Exports and Imports (% of GDP) (V9), FDI Net Inflows (% of GDP) (V10), and Internet Use Rate (V11). Some potentially effective variables such as tertiary enrollment ratio, public expenditure on education (% of GDP), poverty headcount ratio at $3.10 per day, income share held by the highest 10%, income share held by the lowest 10%, and military expenditure (% of GDP) were excluded due to missing observations.

To consider standardized data, cluster analysis was employed. In the light of the findings from the dendogram from hierarchical clustering, we determined four clusters for K-means cluster analysis. It should be noted that the clusters containing countries are different from the four HDI categories (Appendix-I), particularly between the low and medium HDI groups. To determine the main characteristics of clusters, the mean scores displayed in Table 2 were used.

Table 2. Mean Scores of Indicators

Variables	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Maternal Mortality Rate	1.388	-0,26851	-0,635	-0,682
Adolescent Birth Rate	1.222	-0,03044	-0,846	-0,759
Share of Seats in Parliament (Women)	0.1048	-0,28347	-0,4915	0,4091
Female labor force participation rate	0.6726	-0,47464	-0.4661	0.0503
Deaths due to Tuberculosis	1.031	-0,18875	-0,5157	-0,506
Incrase rate of Physicians	-1.034	-0,32569	0,606	0.895
Public Health expenditure (% of GDP)	-0,255	-0,32016	-0,696	0,8337
CO2 Emissions per capita	-0,661	-0,35962	1.945	0,3896
Exports and Imports (% of GDP)	-0.2634	-0,22158	1.456	0,0327
FDI inflows (% of GDP)	0.2238	-0,06918	0,7678	-0.3564
Internet Use rate	-1.115	-0,30259	0,9781	1.108

As presented in Table 2, countries included in Cluster 1 have very low levels of Internet use rate and increase in rate of physicians. However, maternal mortality rate, adolescent birth rate, and deaths due to tuberculosis are at the highest levels. CO2 emissions per capita and percentage of exports and imports are at highest levels for Cluster3. Cluster 4 has the highest level of Internet use rate. It should be noted that the countries in the obtained clusters are different from the four HDI categories. As can be seen in Table 1 in the Appendix, the countries in Cluster 1 mainly correspond to the countries in the low HDI category, and countries in Cluster 4 mostly overlap with countries in the very high HDI category. However, Middle East and North Africa (MENA) countries included in the very high HDI category are not included in Cluster 4; instead, these countries constitute Cluster 3. Most of the countries in the high HDI category and some in the medium HDI country (e.g. South Africa, The Philippines, Indonesia, and India) are in Cluster 2. Although the main characteristics of these clusters are similar to the relevant HDI country categories, the same countries were not obtained with cluster analysis using the 11 indicators.

To provide insights about the effectiveness of indicators used to determine the classification of countries, MDS was employed. The visual representation considering the 11 indicators is presented in Figure 2. A stress value of 0.08 was obtained, which is an evidence of a “good” fit between the original distances and predicted distances on a two-dimensional plot. As can be seen from the positions of indicators, maternal mortality rate (V1) and exports and imports as percentage GDP (V9) are unrelated whereas variables V5, V6, V7, V8, and V10 have similar effects.

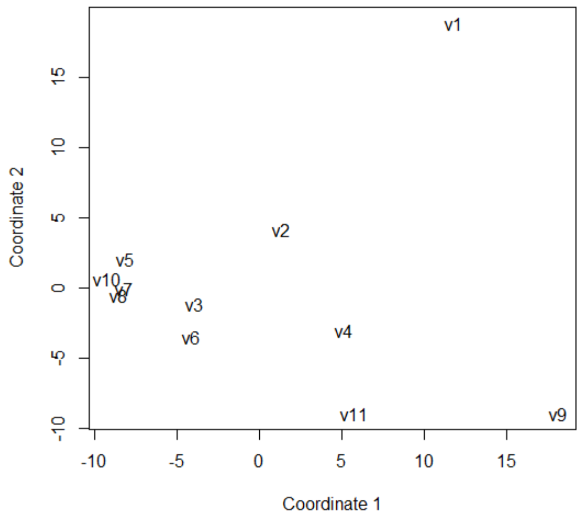


Figure 2. Perceptual Plot of Indicators

To assess the effects of indicators on country categories according to HDI and make a classification, discriminant analysis was employed. However, the assumptions of multivariate normality and homoscedasticity were not provided. [†] Therefore, multinomial logistic regression was used to investigate the effects of the indicators on HDI. Multinomial logistic regression is robust for revealing violations of assumptions of multivariate normality and covariance equality of groups. Furthermore, multinomial logistic regression does not assume a linear relationship between the dependent and independent variables (Akinci et al., 2007).

To determine the difference between the model without independent variables and the model with independent variables, a likelihood ratio Chi-Square test was performed and revealed that at least one of the regression coefficients in the model is not equal to zero (p=0.000). McFadden's pseudo R square for the model was 0.814. Since multinomial logistic model estimated k-1 [‡] equations, three equations are displayed in Table 3 with the high HDI country group chosen as the base outcome.

The first part of Table 3 displays indicators associated with very high scoring HDI countries. The rate of Internet use and percentage of exports and imports of GDP had significant effects for very high HDI countries relative to high HDI countries. With a one-unit increase in the indicator of percentage of exports and imports, the multinomial log odds for a country having a very high HDI relative to high HDI is expected to decrease 0.016 units while all the variables in the model remain constant. Considering the RRR statistic,[§] it can be interpreted that with a one-unit increase of the percentage of exports and imports variable, the relative risk of being in very high HDI country group is expected to decrease by a factor of 0.98 when the other variables are held constant. For the rate of Internet use variable, a one-unit increase is expected to increase the multinomial log odds by 1.38 units for the very high HDI category relative to the high HDI group, which means that with an increase of one unit in the rate of Internet use, the relative risk of being a very high HDI country is 1.38 times more likely. The second (middle) part of Table 3 corresponds to indicators associated with countries with medium HDI scores. Deaths due to tuberculosis, increase in rate of physicians, CO_2 emissions per capita, FDI net inflows as percentage of GDP, and Internet use rate were found to be significant variables that have effects on medium HDI countries relative to high HDI countries. With a one-unit increase of deaths due to tuberculosis, the multinomial log odds for medium HDI countries relative to high HDI countries is expected to increase 0.22 while the other variables in the model remain constant. Similarly, with a one-unit increase in increase in rate of physicians, the multinomial logs odds for medium HDI countries relative to high HDI countries is expected to increase by 0.16 units.

However, with separate one-unit increases of CO_2 emissions per capita, FDI net inflows as percentage of GDP, and Internet use rate, the relative risk of being in the medium HDI group relative to the high HDI group is expected to decrease by factors of 0.51, 0.76, and 0.89, respectively. The last part of Table 3 refers to the indicators associated with countries with low HDI scores. The significant variables and their effects on the model for medium HDI countries relative to high HDI countries are similar for the model with low HDI countries relative to high HDI countries. However, there is another significant indicator, namely, maternal mortality rate. A one-unit increase in the maternal mortality rate increases the relative risk of being in the low HDI group by a factor of 1.05.

Therefore, on the basis of the 11 socio-economic and health indicators, countries were classified into four groups. According to cluster means (Table 2), health related indicators (maternal mortality rate, deaths due to tuberculosis, and increase in rate of physicians) are effective for the classification of countries. Countries with high mean scores for health indicators are characteristically countries in the low HDI category. Alongside the health indicators, rate of Internet use, exports and imports as percentage of GDP, and CO_2 emissions per capita are also distinctive features for classification. Countries with high mean scores for the aforementioned indicators are characteristically countries in the high and very HDI categories. The perceptual map of indicators obtained from MDS analysis (Figure 2) presents the different patterns of maternal mortality rate from the remaining indicators. The findings of multinomial logistic regression (Table 3) indicate that maternal mortality rate and deaths due to tuberculosis are distinctive features of low HDI countries relative to high HDI countries. Rate of Internet use is a striking indicator that separates very high HDI countries from the remaining countries. For medium and low HDI countries relative to high HDI countries, variables of FDI inflow and CO_2 emissions per capita are decisive. Economic indicators, namely, exports and imports as percentage of GDP are important for classifying very high HDI countries relative to high HDI countries.

† Multivariate normality was examined with the Hz test and Royston test. For the homogeneity of covariances, Box-M tests were used.

‡ k is the number of levels of the dependent variable.

§ RRR refers to the relative risk ratio, which is calculated by exponentiating the multinomial logit coefficients.

Table 3. Parameter Estimates of Multinomial Logistic Regression

	Very High HDI				***Medium HDI***				***Low HDI***			
Base: High HDI	**Coef.**	**+Robust Std.Err.**	**Sig.**	**RRR***	**Coef.**	**Robust Std.Err.**	**Sig.**	**RRR***	**Coef.**	**+Robust Std.Err.**	**Sig.**	**RRR***
Maternal Mortality Rate	-0,0489	0,0517	0,3440	0,9522	0,0336	0,0221	0,1280	1,0342	0,0506	0,0238	0,03**	1,0519
Adolescent Birth Rate	-0,0417	0,0290	0,1510	0,9592	0,0483	0,0299	0,1060	1,0494	0,0497	0,0367	0,1770	1,0509
Share of seats in Parliament	0,1626	0,1474	0,2700	1,1765	-0,0192	0,0443	0,6640	0,9809	-0,0352	0,0760	0,6430	0,9654
Female labour forceparticipation rate	0,0098	0,0567	0,8620	1,0098	-0,0415	0,0429	0,3330	0,9594	-0,1023	0,0702	0,109	0,9028
Deaths due to Tuberculosis	-0,3995	0,2747	0,1460	0,6706	0,2247	0,1094	0,04**	1,2520	0,2488	0,1170	0,033**	1,2825
Physicians	-0,0338	0,4458	0,4480	0,9667	0,1647	0,0699	0,019**	1,1791	-0,1055	0,2969	0,7220	0,8999
Public health expenditure % of GDP	0,1251	0,3199	0,6960	1,1332	-0,4484	0,2939	0,1270	0,6386	0,1836	0,5381	0,7330	1,2015
CO2 emissions per capita	0,0102	0,0728	0,8080	1,0102	-0,6721	0,2973	0,02**	0,5106	-1,7090	0,8990	0,058***	0,1811
Exports and Imports % of GDP	-0,0159	0,0090	0,076***	0,9843	0,0363	0,0223	0,1030	1,0369	-0,0082	0,0492	0,8680	0,9918
FDI, inflows % of GDP	-0,1121	0,0699	0,1090	0,8940	-0,2703	0,1350	0,045**	0,7632	-0,3113	0,1492	0,037**	0,7325
Internet users rate	0,3245	0,1358	0,017**	1,3834	-0,1166	0,0462	0,012**	0,8899	-0,1932	0,0569	0,001*	0,8243
Intercept	-20,6781	8,4521	0,0140	-	-0,2052	1,9233	0,9150	-	1,4901	3,7356	0,6920	-

+ : Robust Standardize Errors
*, **, *** represent significance at 1, 5 and 10 % levels, respectively.

5. CONCLUSIONS

The HDI index has become one of the most widely used methods of ranking the development of countries according standards of living, life expectancy and literacy levels. Notwithstanding, the accuracy of the HDI has come under increased criticism in the literature due to a lack of representativeness and gaps in the data (Srinivasan, 1994; Ranis et al., 2006; Wolf et al., 2010).

The present study considers 11 socioeconomic and health developmental indicators for 175 countries, classifying them and assessing the effects of these indicators on the HDI country categories. We deduced four country clusters using cluster analysis, and used these country clusters as the basis of our HDI country categories. Comparison were made between the countries in the HDI categories and those in the clusters to reveal a number of startling differences. However, in evaluating the means scores for the clusters, we also revealed a number of commonalities with HDI categories. For instance, the rate of Internet use, exports and imports as percentage of GDP and CO_2 emissions per capita distinguish developed and undeveloped countries, with these indicators having higher mean scores for high and very high HDI countries. On the other hand, cluster mean scores for maternal mortality rate, deaths due to tuberculosis and increase in rate of physicians were greater for undeveloped countries, which were largely consistent with the low HDI countries. Notwithstanding, MDS analysis revealed that the maternal mortality rate was distinct from other indicators, as indicated by its position on the perceptual map.

Multinomial analysis was used to investigate the effect of the indicators on the HDI country categories; the findings alluding to similar interpretations as what were found with cluster analysis. Maternal mortality rate and deaths due to tuberculosis were the distinguishing features of low HDI countries relative to high HDI countries. Regarding the indicators of FDI inflow and CO_2 emissions per capita, these are more distinctive for medium and low HDI countries relative to high HDI countries. Specifically, if one of these indicators increases by just one unit, a country is much more likely to be in the high HDI category relative to the medium and low HDI categories. For classification in the very high HDI category relative to the high HDI category, exports and imports as a percentage of GDP is prominent. Specifically, a one unit increase in this indicator means that a country is much more likely to be in the high HDI group.

Additionally, the findings of this study reveal shortcomings in the classification of low income countries based on health issues and women's involvement in parliament and work. These turned out not to be particularly effective indicators for very high, medium, and low HDI countries relative to high HDI countries.

References

Agresti A., (2002). Categorical Data Analysis,John Wiley &Sons,New Jersey.

Akinci S., Kaynak E., Atilgan E.,Aksoy S.(2007). Where does the logistic regression analysis stand in marketing literature?A comparison of the Market Positioning of Prominent Journals, European Journal of Marketing. 41(5/6),537-567.

Antony GM, Visweswara Rao K, Balakrishna N.,(2001). Human Development Index: It's Suitability for the Assessment of Health and Nutritional Status of Indian States, EPW, Bombay, India, 36 ,p. 2976-9.

Barreiro I.C, (2006). Human Development assessment through the Human-Scale Development approach, Ph-D Thesis, Universitat Politecnica De Catalunya.

Bedir S., Yilmaz,V.M., (2015). CO2 Emissions and human development in OECD countries: granger causality analysis with a panel data approach, Eurasian Economic Review,p:1-14,doi 10.1007/s40822-015-0037-2

Boutayeb A., Serghini M.(2006),Heath indicators and human development in the Arab region, Int. J. Health Geogr,5(61), doi: 10.1186/1476-072X-5-61

Costa L., Rybski D., Kropp J.P., (2011). A Human Development Framework for CO2 Reductions. PLoS ONE 6(12): e29262. doi:10.1371/journal.pone.0029262

Fukuda-Parr, S.,(2001). Indicators of Human Development and Human Rights-Overlaps, Differences… and What about the Human Development Index? Statistical Journal of the United Nations Economic Commission for Europe Vol. 18., 239-248

Hintze,NCSS User's Guide I,NCSS, 2007 http://www.ncss.com/wpcontent/uploads/2012/09/NCSSUG1.pdf

Ledoleter, J. (2013),Data Mining and Business Analytics with R, John Wiley &Sons, New Jersey.

Lee K., Park, S., Knoshnood B., Hsieh KL., Mittendorf R., (1997).Human development index as a predictor of infant and maternal mortality rates. J.Pediatr., 131(3), 430-3.

Mooi E., and Sarstedt M., (2011).A Coincise Guide to Market Research, Springer-Verlag Berlin Heidelberg.

Njoh A.J., (2003). Urbanization and development in sub-Saharan Africa. Cities 20: 167–174

Öztürk, S.G., (2007). Classifying and predicting country types through development factors that influence economic, social, educational and health environments of countries,SWDI Proceedings papers S759, p.665-674.

Ranis, Stewart, Samman, (2006). Human Development: Beyond the Human Development Index, Journal of Human Development,7(3), p. 322-358.

Reiter S.L., Steensma H.K., (2010).Human Development and Foreign Direct Investment in Developing Counties: The Influence of FDI Policy and Corruption. World Development, 38(12), 1678-1691.

Saito M., (2003). Amartya Sen's Capability Approach to Education: A Critical Exploration. Journal of Philosophy of Education, 37: 17–33. doi: 10.1111/1467-9752.3701002

Self S., Grabowski R., (2003). How Effective is Public Health Expenditure in Improving Overall Health?A Cross-country Analysis, Applied Economics, 35,835-845.

Sen A., (1992). Inequality Reexamined, Russell Sage Foundation, New York.

Sharma B.,Gani A., (2004). "The Effects of Foreign Direct Investment on Human Development," Global Economy Journal, De Gruyter, 4(2), p. 1-20,

Sharma, S. D.,(1997). Making the Human Development Index (HDI) gender-sensitive. Gender and Development Vol. 5, No 1., 60—61

Soubbotina T.P., Sheram K.,(2000). Beyond Economic Growth, The World Bank.

Srinivasan T.N., (1994). Human Development: A New Paradigm or Reinvention of the Wheel?, American Economic Review, 84(2), p. 238-243.

Stanton E. ,(2007). The Human Development Index: A history,working paper,University of Massachussetts Amherst, http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1101&context=peri_workingpapers

Tan P.N., Steinbach M., Kumar V., (2005).Introduction to Data Mining, Pearson, New York

United Nations Development Programme (UNDP), Human Development Report (1990).

United Nations Development Programme (UNDP), Human Development Report (2014).

Wolff H., Chong H. and Auffhammer M., (2010). Classification, detection and consequences of data error: evidence from the human development index',

http://faculty.washington.edu/hgwolff/EJOnlineWebVersionofHDI_Wolffetal2010.pdf

Appendix

Country Classifications based on 11 indicators

Countries	HDI	Cluster	Countries	HDI	Cluster
Norway	1	4,000	Tunisia	2	2,000
Australia	1	4,000	Colombia	2	2,000
Switzerland	1	4,000	Jamaica	2	2,000
Denmark	1	4,000	Tonga	2	2,000
Netherlands	1	4,000	Belize	2	2,000
Germany	1	4,000	Maldives	2	3,000
Ireland	1	3,000	Samoa	2	2,000
United States	1	4,000	Botswana	3	2,000
Canada	1	4,000	Moldova (Republic of)	3	4,000
New Zealand	1	4,000	Egypt	3	2,000
Singapore	1	3,000	Turkmenistan	3	3,000
Hong Kong	1		Gabon	3	1,000
Sweden	1	4,000	Indonesia	3	2,000
United Kingdom	1	4,000	Paraguay	3	2,000
Iceland	1	4,000	Uzbekistan	3	2,000
Korea (Republic of)	1	4,000	Philippines	3	2,000
Israel	1	4,000	El Salvador	3	2,000
Luxembourg	1	3,000	South Africa	3	2,000
Japan	1	4,000	Viet Nam	3	2,000
Belgium	1	4,000	Bolivia	3	1,000
France	1	4,000	Kyrgyzstan	3	2,000
Austria	1	4,000	Iraq	3	2,000
Finland	1	4,000	Cabo Verde	3	2,000
Slovenia	1	4,000	Guyana	3	2,000
Spain	1	4,000	Nicaragua	3	2,000
Italy	1	4,000	Morocco	3	2,000
Czech Republic	1	4,000	Namibia	3	2,000
Greece	1	4,000	Guatemala	3	2,000
Estonia	1	3,000	Tajikistan	3	2,000
Brunei Darussalam	1		India	3	2,000
Cyprus	1	4,000	Honduras	3	2,000
Qatar	1	3,000	Bhutan	3	2,000
Slovakia	1	4,000	Timor-Leste	3	1,000
Poland	1	4,000	Syrian Arab Republic	3	2,000
Lithuania	1	4,000	Vanuatu	3	2,000
Malta	1	4,000	Congo	3	
Saudi Arabia	1	3,000	Kiribati	3	
Argentina	1	4,000	Equatorial Guinea	3	1,000
United Arab Emirates	1	3,000	Zambia	3	1,000
Chile	1	2,000	Ghana	3	1,000
Portugal	1	4,000	Bangladesh	3	1,000
Hungary	1	4,000	Cambodia	3	1,000
Bahrain	1	3,000	Sao Tome and Principe	3	2,000
Latvia	1	4,000	Kenya	4	1,000
Croatia	1	4,000	Nepal	4	1,000
Kuwait	1	3,000	Pakistan	4	2,000
Montenegro	1	2,000	Myanmar	4	
Belarus	2	4,000	Angola	4	1,000
Russian Federation	2	4,000	Swaziland	4	1,000
Oman	2	3,000	Tanzania	4	1,000
Romania	2	2,000	Nigeria	4	1,000
Uruguay	2	4,000	Cameroon	4	1,000
Kazakhstan	2	4,000	Madagascar	4	1,000
Barbados	2	4,000	Zimbabwe	4	

Countries	HDI	Cluster	Countries	HDI	Cluster
Bulgaria	2	4,000	Mauritania	4	1,000
Panama	2	2,000	Solomon Islands	4	2,000
Malaysia	2	3,000	Papua New Guinea	4	
Mauritius	2	2,000	Comoros	4	2,000
Seychelles	2		Yemen	4	2,000
Trinidad and Tobago	2	3,000	Lesotho	4	1,000
Serbia	2	4,000	Togo	4	1,000
Cuba	2		Haiti	4	
Lebanon	2	3,000	Rwanda	4	1,000
Costa Rica	2	2,000	Uganda	4	1,000
Iran	2	2,000	Benin	4	1,000
Venezuela	2	2,000	Sudan	4	2,000
Turkey	2	2,000	Djibouti	4	1,000
Sri Lanka	2	2,000	South Sudan	4	
Mexico	2	2,000	Senegal	4	1,000
Brazil	2	2,000	Afghanistan	4	1,000
Georgia	2	4,000	Côte d'Ivoire	4	1,000
Saint Kitts and Nevis	2		Malawi	4	1,000
Azerbaijan	2	4,000	Ethiopia	4	1,000
Grenada	2		Gambia	4	1,000
Jordan	2	2,000	Congo	4	1,000
Macedonia	2	4,000	Liberia	4	1,000
Ukraine	2	4,000	Guinea-Bissau	4	
Algeria	2	2,000	Mali	4	1,000
Peru	2	2,000	Mozambique	4	1,000
Albania	2	2,000	Sierra Leone	4	1,000
Armenia	2	2,000	Guinea	4	1,000
Bosnia and Her.	2	4,000	Burkina Faso	4	1,000
Ecuador	2	2,000	Burundi	4	1,000
China	2	2,000	Chad	4	1,000
Fiji	2	2,000	Eritrea	4	1,000
Thailand	2	2,000	Central African Republic	4	1,000
Dominica	2		Niger	4	1,000
Libya	2	2,000			

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2015.04.01.OR.01

OPTIMAL CITY STREET NETWORK DESIGN UNDER UNCERTAINTY

Andrew HOWE*

KAPSARC, Riyadh, Saudi Arabia (Affiliated after article accepted)

Risk Dynamics Consultancy, Istanbul, Turkey

Received: 22 January 2016

Accepted: 14 March 2016

Abstract

In this article, we address the problem of city street network design - specifically that of using one-way vs. two-way streets - from a different perspective than existing published literature. While at least one author acknowledges that motorist familiarity should be a factor in network design, this has not been empirically modeled. Instead of implicitly assuming motorists travel optimal paths, we explicitly model motorist unfamiliarity and uncertainty with an area. Furthermore, while the published research uses VMT or similar metrics to evaluate network design, we propose the ratio of actual VMT to optimal VMT as a more appropriate metric, with a target ratio of unity. We develop two simple idealized city street grids: one in which all streets are two-way, and a second of perfectly alternating one-way streets. Motorists are simulated traveling on both grids while varying the level of unfamiliarity and uncertainty. For each motorist, the ratio of actual to optimal VMT is measured and recorded. Our results suggest that travel efficiency for visiting motorists unfamiliar with an area will always be highest for one-way street networks. The policy this suggests is that one-way network city street designs should be preferred when there are likely to be a high proportion of motorists who are unfamiliar with the area. This conclusion goes against the prevailing wisdom, since most analysis evaluates network designs based on minimizing VMT, assuming motorists travel optimal paths.

Keywords: Transportation Modeling, Transportation Network Design, Monte Carlo Simulation

BELİRSİZLİK ALTINDA OPTİMAL ŞEHİR SOKAK AĞI TASARIMI

Özet

Bu çalışmada, mevcut literatürden farklı bir perspektif ile özellikle tek yönlü sokakların çift yönlü sokaklarla karşılaştırılmalı olarak kullanıldığı şehir-sokak ağ tasarımı problemi ele alınmıştır. Literatürde sürücünün bölgeye aşinalığının ağ tasarımında bir faktör olması gerektiğini ifade eden en az bir çalışma mevcuttur ancak bu sonuç deneysel olarak modellenmemiştir. Bu çalışmada motorlu taşıt kullanıcılarının dolaylı yollar yerine en uygun yollarda seyahat ettiği varsayımı altında, motorlu taşıt kullanıcısının bir bölgeye aşina olmaması ve belirsizlik açık bir şekilde modellenmiştir. Buna ek olarak, önceki çalışmada ağ tasarımını değerlendirmek için VMT (bir aracın belirli bir bölgede belirli bir zamanda katettiği mesafenin mil cinsinden ifadesi) ya da benzer ölçüm birimleri kullanılmışken, bu çalışmada daha uygun bir ölçüm aracı olarak, gerçekleşen VMT ile optimal VMT'nin oranlandığı bir ölçüm birimi önerilmiştir. Araştırmada tüm sokakların çift yönlü olduğu ve alternatif tek yönlü olduğu iki adet basit şehir ve sokak çizelgesi geliştirilmiştir. Sürücüler için değişen düzeylerde bölgeye aşina olmama ve belirsizlik durumları iki çizelge için de simüle edilmiştir. Gerçekleşen VMT ile optimal VMT'nin oranı her bir sürücü için ölçülerek ve kaydedilmiştir. Analiz sonuçlarına göre bir bölgeye aşinalığı olmayan sürücüler için tek yönlü şehir ağının her zaman daha verimli olacağı bulgusu elde edilmiştir. Bu doğrultuda bir bölge ile aşinalığı olmayan sürücülerin yoğunlukta olduğu durumlarda tek yönlü şehir ağının tercih edilmesi gerektiği önerilmiştir. Sürücülerin en uygun yollarda seyahat ettiği varsayımı altında, çoğu analiz, ağ tasarımını VMT'yi minimize etme temelinde değerlendirdiğinden, bu sonuç yaygın görüşe farklı bir bakış açısı getirmektedir.

Anahtar Kelimeler : Ulaştırma Modeli, Ulaştırma Ağ Tasarımı, Monte Carlo Simülasyonu

* ahowe42@gmail.com (Corresponding author)

1. INTRODUCTION

In Walker et al. [3], the authors intimate:

One-way streets do not pose a major inconvenience for commuters and regular visitors to the downtown; these motorists have learned the downtown network and know the " best route" to their destination. Rather, it is the occasional visitors to downtown who are often confused and disoriented on encountering a one-way street network.

Many of us drivers have been stymied by this, and I have often cursed finding the direction I want to turn denied by a one-way street. It is understood that driving in a city that has one-way streets is just as simple as if they are all two-way streets *when the driver knows where he is going* [†]. However, what about these visitors? We are unaware of any published research evaluating this " visitor effect" of path uncertainty on the design of city street networks.

The literature on city street network design, as it relates to one-way vs. two-way streets is rather limited and consistently focuses on different types of analyses than we do in this article. Much of it is at least partially based on the knowledge that "*One-way links, although lengthening some shortest paths, are faster, effectively reducing the travel time in the permitted direction*."[1] In Drezner and Wesolowsky [1], the authors explicitly account for the speed / distance characteristics of both one-way and two-way streets to optimize a transportation network.

In [3], Walker *et al*. explored the apparently contentious issue of converting downtown city streets from one-way to two-way. Their analysis considered multiple facets, such as street capacity, and motorists' trip characteristics. External factors considered were pedestrians' interactions with drivers, public transit, and economic retail. The "measures-of-goodness" related to motorists they used included number of turns and the ubiquitous vehicle miles of travel (VMT). In [4], the authors compared the difference in emissions over one-way and two-way street networks, based again on VMT. They claim their analysis demonstrates that two-way street networks lower emissions over one-way networks, depending on traffic demand.

The final, and most interesting article detailed here, is Gayah and Daganzo [2]. In this article, the authors challenge the belief held by traffic engineers and city planners that one-way street networks handle traffic more efficiently. They developed an idealized network flow model with either one-way or two-way streets and measured the trip-serving capacity. Their results suggest that two-way streets are more efficient, especially for short trips.

The surveyed authors have focused on various measures of city street network design, but most seem to have completely overlooked the fact that motorists do not always follow optimal paths. For any network, there exists an optimal path (possibly many) between any two points, A and B. Adding restrictions to any network will generally increase the length of the optimal path averaged over all origins and destinations; this is so certain as to be nearly axiomatic. Hence, the VMT for a network of two-way streets should necessarily be less (or equal) to the VMT if there were one-way streets. We suggest a more appropriate metric is the ratio of actual VMT to optimal VMT. When a path followed is the optimal, this ratio will be ; a value greater than unity indicates a less efficient path.

To this end, we have developed a city street network model that allows us to simulate traffic flow on either a grid of one-way streets, or a grid of two-way streets. With both networks, we randomly select starting points and use Monte Carlo simulation to simulate motorists driving to a destination, varying the amount of uncertainty and unfamiliarity. For each driver, we measure the ratio of actual to optimal VMT.

Our results suggest that as unfamiliarity increases, this extra-distance ratio increases log-linearly when the streets are all one-way. However, when a network of two-way streets is used, it increases at a higher rate. The variability in the extra-distance ratios shows the same pattern. Our results show that an idealized one-way street network design is basically never worse than a completely unrestricted design.

In Section 2, we describe the city street network model and it's capabilities. We also detail the Monte Carlo simulation used to assess the joint impact of unfamiliarity and street network design on VMT. We describe and interpret the results of our simulations in Section 3, then finish up with some concluding remarks in the last section.

2. THE CITY STREET NETWORK SIMULATION MODEL

2.1. The Model

We begin with a city street network that is a simple grid as shown in the left pane of Figure 1. We use a secondary grid to encode the allowed directions of travel in each block for two possible idealized city street networks:

[†] To the best of our knowledge, however, this has not been empirically demonstrated

- all streets are two-way, with no restrictions
- all streets are one-way, with the direction of travel for parallel streets alternating perfectly

These are shown, respectively in the left and right panes (respectively) of *Figure 2*.

Directions are coded such that U=up, D=down, L=left, and R=right; in *Figure 1*, some of the end blocks are blocked off, since the allowed direction of travel would only lead off the grid. In the grids, we indicate a randomly chosen destination with the highlighted **B**.

The network model is completely malleable; it can be modified to allow any possible travel direction restrictions, and blockages of any shape can be placed anywhere in the network. An example of a complex network is shown in *Figure 1*. After designing the network and specifying the allowed directions, the model automatically checks itself for consistency by verifying that every **A** can reach the specified **B**. It can also automatically identify and record in a tertiatry grid the optimal direction and distance from every **A** to the target **B**. When there are travel restrictions, however, these optima must be manually verified and edited.

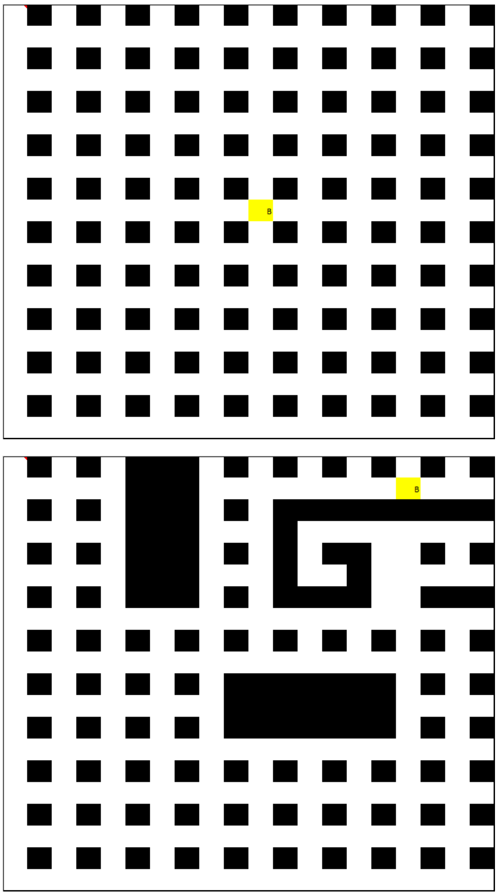


Figure 1. City Street Network Model Grids. First one is a Simple Grid, second is a Sample Complex Grid.

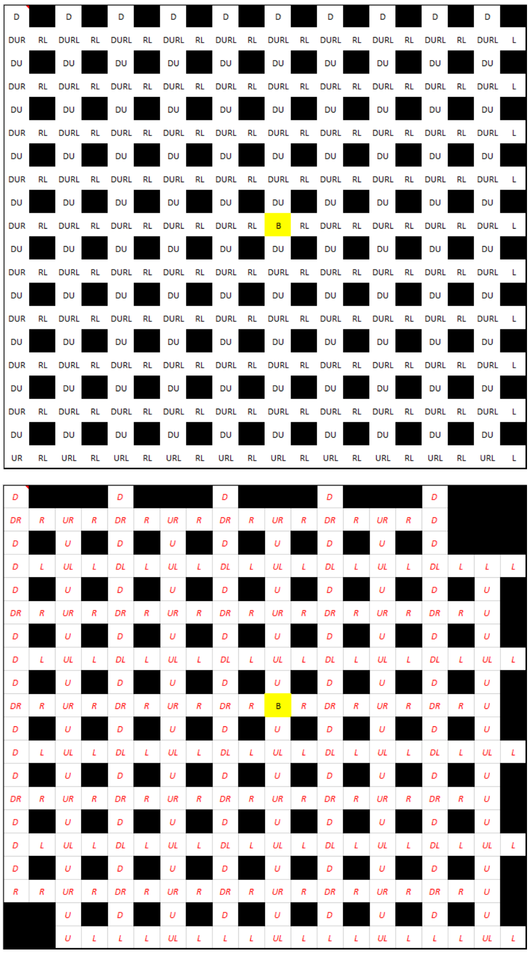


Figure 2. City Street Network Models - Allowed Directions. First Pane is Network O, Second Pane is Network R.

2.2. Simulation

Our Monte Carlo simulation proceeds in two stages. First, an available block is randomly selected as the origin. Secondly, we simulate a motorist following a sequence of paths to drive from **A** to **B**. The complete algorithm is shown here:

I. Randomly select origin **A** uniformly from available blocks

II. Driver begins at A

i. Randomly select the driver's unfamiliarity for this path from three choices: Informed, Uninformed, Completely Wrong using a discrete probability distribution. Probabilities used are specified by the modeler. The resultant travel for this path is generated according to these rules:

- *Informed:* driver travels optimal direction and distance
- *Uninformed:* direction is chosen uniformly from allowed directions, and distance is simulated using a Poisson distribution covering up to the length of the grid, but with a mean of four blocks, and 90% confidence interval of between one and eight blocks
- *Completely Wrong:* driver travels the worst direction until forced to turn

ii. The driver's uncertainty is simulated using a Triangular distribution with minimum of 80%, maximum of 120%, and mode of 100%; this is used to scale the distance traveled generated in *step (i)*

iii. The motorist drives along the path specified by *(i)* and *(ii);* if he has arrived at **B** , go to *III*, else return to *(i)*

III. Record the VMT and number of paths required, compute their ratios to the optimal quantities

IV. If still need to simulate more drivers per **A** , return to II; if number of drivers required is met but still need to simulate more **A**'s, return to *I*; otherwise exit

The simulation parameters which can be set by the user are:

- Number of origins from which to simulate travel
- Number of drivers to simulate driving from selected **A**
- Relative frequencies of unfamiliarity (Informed, Uninformed, Completely Wrong)
- Range of uncertainties for Triangular distribution

The entire city street network simulation model is developed in Microsoft Excel and Palisade's *@RISK*.

3. NUMERICAL RESULTS

For all our results, we model the uncertainty with the default parameters of 80% and 120% for the Triangular distribution. When simulating the drivers' unfamiliarity, we keep the probability of being Completely Wrong at a low 5% to simulate wrong turns, and vary the probability of being Informed between 50% and 90% in increments of 10%. The probability associated with being Completely Wrong is simply the remainder. For each of the ten scenarios (five levels of unfamiliarity for two grids), we generated 100 random origins, and simulated 50 motorists traveling from A to B .

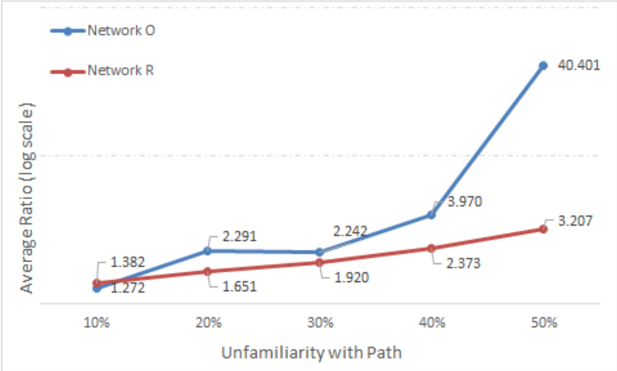


Figure 3. Behavior of Ratio of Actual to Optimal VMT as Unfamiliarity Increases (Averages).

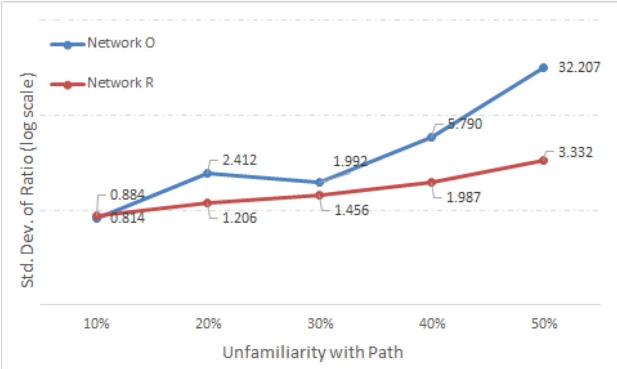


Figure 4. Behavior of Ratio of Actual to Optimal VMT as Unfamiliarity Increases (Standard Deviations).

We first analyzed the affect of motorists' unfamiliarity on the ratio of actual to optimal number of turns, but found no consistent difference between the two network designs. When analyzing the actual to optimal VMT ratio, however, a clear pattern emerged.

Our results suggest that as unfamiliarity increases, the ratio of actual VMT to optimal VMT increases log-linearly under city street network design R. However, when a network of two-way streets is used, it increases at a higher rate. The variability in the extra-distance ratios shows the same patterns. These ratios are plotted in *Figure 3* and *Figure 4*, respectively. Our results empirically support the conventional wisdom that when motorists are very familiar with an area, whether the streets are a one-way or two-way network, they travel close to the optimal path. The extra distance driven over the optimal is relatively low for both network designs.

More importantly, our results show that when motorist familiarity is explicitly modeled, street network design R is basically never less efficient than a network design based on two-way streets. While occasional visitors may become disoriented, they actually end up driving less extra distance over the optimum than if a two-way street network was in place. For example, with 80% familiarity, average VMT may be more than twice as much as optimal when there are two-way streets, while it would only be 65% higher if there were one-way streets. In the case of the two-way network, the average optimal VMT was 10 blocks so the average actual VMT (for 80%) would be 23 blocks. The same quantities for the one-way network design were 12 and 20 blocks, respectively. Hence, the visiting motorist actually drivers further on average, when the streets are all two-way.

4. CONCLUDING REMARKS

In this article, we have addressed the problem of city street network design from a different perspective than existing published literature. Instead of implicitly assuming motorists travel optimal paths, we explicitly modeled motorist unfamiliarity and uncertainty with an area. We developed two simple idealized city street grids: one in which all streets are two-way, and a second of perfectly alternating one-way streets. Motorists are simulated traveling on both grids while varying the level of unfamiliarity and uncertainty. For each motorist, the ratio of actual to optimal VMT is measured and recorded.

Our results suggest that travel efficiency for visiting motorists unfamiliar with an area will always be highest for one-way street networks. The policy this suggests is that one-way network city street designs should be preferred when there are likely to be a high proportion of motorists who are unfamiliar with the area. This conclusion goes against the prevailing wisdom, since most analysis evaluates network designs based on minimizing VMT, assuming motorists travel optimal paths.

References

[1] Zvi Drezner and George O. Wesolowsky. Selecting an Optimum Configuration of One-Way and Two-Way Routes. Transportation Science, 31(4):386–394, November 1997.

[2] Vikash V. Gayah and Carlos F. Daganzo. Analytical Capacity Comparison of One-Way and Two-Way Signalized Street Networks. Transportation Research Record, 2301:76–85, December 2012.

[3] G. Wade Walker, Walter M. Kulash, and Brian T. McHugh. Downtown Streets: Are We Strangling Ourselves with One-Way Streets? TRB Circular E-C019: Urban Street Symposium, 1999.

[4] Jinghui Wang, Lei Yu, and Fengxiang Qiao. Micro Traffic Simulation Approach to the Evaluation of Vehicle Emissions on One-way vs. Two-way Streets: A Case Study in Houston Downtown. Presented at the 92nd annual meeting of the Transportation Research Board in Washington D.C., January 2013.

This page blank left intentionally

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2016.04.01.OR.02

COMBINED MULTI-CRITERIA DECISION MAKING APPROACH BASED ON MACBETH AND MULTI-MOORA METHODS

Nilsen KUNDAKCI*

Pamukkale University, Department of Business Administration, Denizli, Turkey

Received: 18 February 2016

Accepted: 22 April 2016

Abstract

Various Multi-Criteria Decision Making (MCDM) methods have been developed to support decision making process. The main aim of all MCDM methods is to obtain ranking of the alternatives and select the best one under conflicting criteria. In this paper, a combined MCDM approach is proposed based on MACBETH (Measuring Attractiveness by a Categorical Based Evaluation TecHnique) and MULTI-MOORA (Multi Objective Optimization on the basis of Ratio Analysis) methods. In this combined approach, the weights of the criteria are determined with MACBETH method and then MULTI-MOORA method is used to obtain the final ranking of the alternatives. At the end of the paper, to illustrate the applicability of the proposed approach an application of the automobile selection of a marble company is also given.

Keywords: MCDM, MACBETH, MULTI-MOORA, Automobile selection

MACBETH VE MULTI-MOORA YÖNTEMLERİNE DAYALI BİRLEŞİK ÇOK KRİTERLİ KARAR VERME YAKLAŞIMI

Özet

Karar verme sürecini desteklemek için çeşitli Çok Kriterli Karar Verme (ÇKKV) yöntemleri geliştirilmiştir. ÇKKV yöntemlerinin temel amacı, alternatiflerin çelişen kriterler altında sıralamalarını elde etmek ve içlerinden en iyisini seçmektir. Bu çalışmada, MACBETH ve MULTI-MOORA yöntemlerine dayalı birleşik bir ÇKKV yaklaşımı önerilmiştir. Bu birleşik yaklaşımda, kriterlerin ağırlıkları MACBETH yöntemi ile belirlenirken, alternatiflerin sıralamasının elde edilmesinde MULTI-MOORA yöntemi kullanılmıştır. Çalışmanın sonunda önerilen yaklaşımın uygulanabilirliğini göstermek için bir mermer işletmesinin otomobil seçim uygulamasına yer verilmiştir.

Anahtar Kelimeler : ÇKKV, MACBETH, MULTI-MOORA, Otomobil seçimi

1. INTRODUCTION

Whether in the daily lives of people or professional life of businesses, there are typically multiple conflicting criteria and a lot of alternatives that need to be evaluated in the decision making process. To aid this decision process, different MCDM methods are developed in the literature. The objective of this paper is to propose a combined MCDM approach based on MACBETH and MULTI-MOORA methods for the first time. And this proposed approach is applied to an automobile selection problem of a marble company in Denizli, Turkey.

There are studies in the literature that consider automobile selection with MCDM methods. For instance, Güngör and İşler [1] solved automobile selection problem with AHP method. Terzi et al. [2] proposed a decision making model based on AHP and goal programing for automobile selection. Soba [3] applied PROMETHEE method for selecting the best automobile by considering price, fuel, maximum speed, horse power and performance criteria. Sakthivel et al. [4] proposed hybrid MCDM technique for the selection of the best car by integrating

* nilsenk@pau.edu.tr (Corresponding author)

Fuzzy Analytical Hierarchy Process (FAHP) with Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) method. Yavaş et al. [5], examined the customers' car selection criteria and used AHP & ANP methods to rank these criteria. The main difference of this paper from other studies that consider automobile selection problem is to use a combined approach based on MACBETH and MULTI-MOORA methods for the first time.

This paper is organized as follows: In the second section MACBETH method is explained and then MULTI-MOORA method is introduced in the third section. In the fourth section, application of the combined approach in a marble company is given. Finally, results of the application are discussed and suggestions for future research are given in the fifth section.

2. MACBETH Method

MACBETH (**M**easuring **A**ttractiveness by a **C**ategorical **B**ased **E**valuation **T**ec**H**nique) was firstly proposed by Bana e Costa, Vansnick and De Corte in 1990s. It was applied to different fields after introduced in the XIth International Conference on MCDM [6]. MACBETH is a MCDM method that helps to build a quantitative model of values and it avoids forcing decision makers to produce direct numerical representations of their preferences [7]. It helps the decision makers to rank the alternatives based on aggregated measurement of relative weighted attractiveness of alternatives with respect to several decision criteria [8].

In the literature there are various studies that apply MACBETH method to different fields. For instance it is used to solve complex strategic problems of Santa Catarina textile industry [9], for conflict dissolution in the construction of a new railway [10], to analyze spatial conflicts in the investment policy of new inter-municipal road-links [11], to assign priorities for maintenance, repair and refurbishment in managing a municipal housing stock [12], for strategic town planning [13], to help credit granting decisions in banking sector [14], for bid evaluation processes in public call for tenders [15], to solve career choice problem [7], for a certain model of coalition formation to determine stable governments [16], for multi-criteria industrial performance expressions [17], for the evaluation and comparison of the technical performance of three hydrogen storage technologies [18], to measure the satisfaction degree of services orchestration to the quality attributes requirements [19], to solve real time supplier selection problems [20], to build a multidimensional value-based population health indices [21], to solve facility layout selection problems [8].

While applying MACBETH method the relevant steps adapted as presented below:

Step 1. Firstly decision criteria are defined and then value tree is formed.

Step 2. After forming the value tree, alternatives are determined. Then the ordinal performance levels representing the possible performance of the alternatives with respect to a particular criterion are defined. Minimum two reference levels are required to be identified as upper reference (good) level and lower reference (neutral) level. On MACBETH scale, the upper reference level has a score of 100, while the lower reference level has a score of 0. Here, 100 does not necessarily represent the best possible score and 0 does not denote the worst performance of an alternative for a given criterion [20].

Step 3. An *m*x*m* matrix is formed for the alternatives where *m* indicates the number of alternatives for that criterion. In this matrix, alternatives are arranged from left to right according to their importance. This is made to quantify the qualitative performance levels and convert quantitative performance levels into MACBETH scale. Also the same procedure is applied for the criteria.

Step 4. Pairwise comparisons are made for the criteria and alternatives based on difference of attractiveness. MACBETH method helps to map the difference of attractiveness using a set of semantic scale having seven categories arranged in descending order of their importance. The equivalent numerical scales and significances of these semantic scales can be seen in Table 1 [20, 7].

Table 1. Semantic scale of MACBETH

Semantic Scale	Equivalent Numerical Scale	Significance
Null	0	Indifference between alternatives
Very Weak	1	An alternative is very weakly attractive over another
Weak	2	An alternative is weakly attractive over another
Moderate	3	An alternative is moderately attractive over another
Strong	4	An alternative is strongly attractive over another
Very Strong	5	An alternative is very strongly attractive over another
Extreme	6	An alternative is extremely attractive over another

Step 5. The judgments provided by the decision maker are checked for consistency. If the provided judgments are found to be inconsistent, M-MACBETH software suggests possible alterations to make the judgments consistent [12].

Step 6. The consistent judgments are transformed into a suitable numerical scale, identified as the MACBETH scale based on linear programming models.

Step 7. Finally, the weighted global scores representing the overall attractiveness of the considered alternatives are computed using an additive aggregation model to rank the alternatives.

For obtaining MACBETH scores of qualitative performance levels, the following procedure is used [8, 19, 20]:

Firstly, decision maker is asked to compare the pairs of alternatives under each criterion. If the decision maker prefers alternative A_i to $A_{i'}$ for a criterion *j*, this is showed as:

$$A_i > A_{i'} \quad (1)$$

Secondly, the decision maker expresses his/her strengths of preference about the alternatives. These strengths of preference are characterized with semantic scale in Table 1. If the decision maker cannot give his/her strengths of preference but only his/her preferences, this is noted by *P*. The decision maker prefers the alternative A_i to $A_{i'}$ with a strength $h \in \{0,1,2,3,4,5,6\}$ for a criterion *j*,

$$A_i \succ^h A_{i'} \quad (2)$$

This is equivalent to:

$$A_i - A_{i'} = h\alpha \quad (3)$$

where α is a coefficient necessary to meet condition A_i and $A_{i'} \in [0,100]$

Consider an example with four alternatives and their preference of importance for the j^{th} criterion are as $A_2 > A_4 > A_1 > A_3$. If $v_j(A_2)$, $v_j(A_4)$, $v_j(A_1)$ and $v_j(A_3)$ are MACBETH scores for A_2, A_4, A_1, A_3 respectively, then $v_j(A_2)=100$, $v_j(A_3)=0$ and $v_j(A_2) > v_j(A_4) > v_j(A_1) > v_j(A_3)$. Then, decision maker expresses his/her strengths of preferences for alternatives using semantic scale in Table 1. These preference strengths of alternatives for j^{th} criterion are given in Table 2.

Table 2. Preference strengths of alternatives for j^{th} criterion

Alternatives	A_2	A_4	A_1	A_3
A_2 (good)	No	Strong	*P*	*P*
A_4		No	Very Weak	*P*
A_1			No	Moderate
A_3 (neutral)				No

From the data in Table 2, these equation systems can be obtained;

$$v_j(A_2) - v_j(A_4) = 4\alpha \quad (4)$$

$$v_j(A_4) - v_j(A_1) = \alpha \quad (5)$$

$$v_j(A_1) - v_j(A_3) = 3\alpha \quad (6)$$

Here, $v_j(A_2)=100$ (good) and $v_j(A_3)=0$ (neutral). After solving equations (4) - (6), solutions are obtained as; $\alpha = 12.5$, $v_j(A_4)=50$ and $v_j(A_1)=37.5$.

By adopting the same procedure, the quantification of alternatives for all the remaining criteria and the criteria weights can be obtained. Then, the converted MACBETH scores for all the performance measures are multiplied by the respective criteria weights and are added together to find out the overall attractiveness scores for the alternatives. The final overall score is obtained using the following additive value model [8, 14]:

$$V(A_i) = \sum_{j=1}^{n} w_j \left(v_j(A_i)\right) \quad (7)$$

$$\sum_{j=1}^{n} w_j = 1, \quad w_j > 0 \quad and \begin{cases} v_j\left(A_i^{good}\right) = 100 \\ v_j\left(A_i^{neutral}\right) = 0 \end{cases} \quad (8)$$

where w_j indicates weight of the j^{th} criterion. The final ranking of the alternatives is determined based on the $V(A_i)$ values. MACBETH method is supported by M-MACBETH software (http://www.mmacbeth.com/en/ downloads.html) developed using algorithm based on linear programming models [20]. In this paper, the weights of the criteria are determined by using M-MACBETH software.

3. MOORA Method

MOORA (**M**ulti **O**bjective and **O**ptimization on the basis of **R**atio **A**nalysis) method was firstly proposed by Brauers and Zavadskas in 2006 [22]. Although MOORA is a relatively new method it has been applied to different areas in the literature. For example, MOORA method is used for privatization in a transition economy [22], determining the ranking of contractor firms [23], road design [24], evaluation of inner climate [25], regional development studies [26, 27], project management [28], parametric optimization of milling process [29], the selection of optimal network [30], determining the critical path in project management [31], determining the popularity of touristic places [32], supply chain strategy selection [33], selecting the best intelligent manufacturing system [34], personnel selection [35, 36], bank branch location selection [37], ranking cloud storage technology firms [38].

MOORA method has three types namely Ratio System, Reference Point Approach and Full Multiplicative Form. Lastly, the final ranking of the alternatives can be obtained with MULTI-MOORA based on dominance theory.

3.1. Ratio System of MOORA

The steps of Ratio System of MOORA can be summarized as follows [22, 23]:

Step 1. Decision matrix X is formed where x_{ij} indicates the value of i^{th} ($i = 1, 2, \ldots, m$) alternative based on j^{th} ($j = 1, 2, \ldots, n$) criterion

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \quad (9)$$

Step 2. Decision matrix is normalized with Eq. (10)

$$x_{ij}^{*} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^{n} x_{ij}^{2}}} \quad (10)$$

Step 3. Weighted normalized decision matrix is formed with the help of Eq. (11). Here w_j *shows* the weight of the j^{th} criterion.

$$v_{ij} = w_j * x_{ij}^{*} \quad (11)$$

Step 4. Final preference (y_i^{*}) values obtained by using Eq. (12). Here $j = 1,2,\ldots,g$ indicates the criteria to be maximized and $j = g+1, g+2, \ldots, n$ shows the criteria to be minimized.

$$y_i^{*} = \sum_{j=1}^{g} v_{ij} - \sum_{j=g+1}^{n} v_{ij} \quad (12)$$

Step 5. Ranking of the alternatives are obtained by ranking y_i^{*} values in descending order.

3.2. The Reference Point Approach of MOORA

In the Reference Point Approach, the first three step will be the same with Ratio System of MOORA. The other steps can be summarized as [23]:

Step 1. Decision matrix X is formed like in Eq. (9)

Step 2. Decision matrix is normalized with Eq. (10)

Step 3. Weighted normalized decision matrix is formed with Eq. (11).

Step 4. Reference points (r_j) are determined for each criterion. While determining reference points, highest values are chosen for maximization criteria, minimum values are chosen for the minimization criteria.

Step 5. The distance between the alternatives and the reference points are calculated and then the best alternative is determined by using the Tchebycheff Min-Max metric given in Eq. (13).

$$\min_i \left\{ \max_j \left(\left| w_j r_j - v_{ij} \right| \right) \right\} \quad (13)$$

According to the Eq. (13), firstly maximum distance values of each alternative to the reference points under all criteria are determined and then from these values minimum one is chosen as the best alternative. Final ranking of the alternatives are obtained by ranking the maximum distance values in increasing order.

3.3. Full Multiplicative Form of MOORA

The steps of Full Multiplicative Form of MOORA can be given as follows [28]:

Step 1. Decision matrix X is formed like in Eq. (9).

Step 2. Multiplicative ranking index U_i for each alternative are determined with Eq. (14).

$$U_i = \frac{A_i}{B_i} \quad (14)$$

Here $A_i = \prod_{j=1}^{g} x_{ij}^{w_j}$, for i^{th} alternative for the criteria to be maximized ($j = 1,2,\ldots,g$) and $B_i = \prod_{j=g+1}^{n} x_{ij}^{w_j}$ for i^{th} alternative for the criteria to be minimized ($j = g+1, g+2, \ldots, n$) [28, 39].

Step 3. Ranking of the alternatives are obtained by ranking U_i values in descending order.

3.4. MULTI-MOORA

MULTI-MOORA method is based on the dominance theory and aims to reach one final ranking from three ranks obtained with Ratio System, Reference Point Approach and Full Multiplicative Form [40].

MULTI-MOORA was firstly proposed by Brauers and Zavadskas in 2010 [28]. And it is used to solve different MCDM problem by various authors. Brauers and Ginevičius [41] tested the economy of the Belgian regions by using MULTI-MOORA. Brauers and Zavadskas [40] proposed MULTI-MOORA to decide on a bank loan for buying property. Brauers et al. [42] used fuzzy MULTI-MOORA for ranking the EU Member States according to their performance. Baležentis et al. [43], used fuzzy MULTI-MOORA for personnel selection. Brauers and Zavadskas [44] tested the robustness of MULTI-MOORA. Datta et al. [45] proposed grey MULTI-MOORA to select industrial robot. Baležentis and Baležentis [46], explained both crisp and fuzzy MULTI-MOORA and summarized the application areas of MOORA and MULTIMOORA. Aksoy [47] evaluated the performances of companies operated by Turkish Coal Enterprises with AHP-based MULTI-MOORA and COPRAS methods. Obayiuwana and Falowo [39], used MULTI-MOORA for wireless network selection.

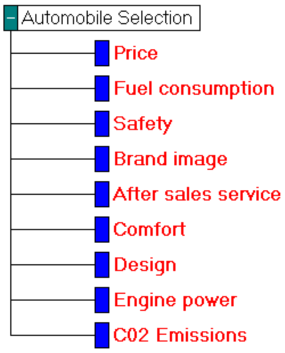
4. APPLICATION

In this part, the applicability of the combined approach is illustrated. For this reason, a MCDM problem of a marble company located in Denizli, Turkey tried to be solved with this approach. A marble company decided to purchase an automobile for the general manager of the company. So it is aimed to select the most appropriate automobile for the general manager as an official car. After a preliminary research they decided to purchase diesel automobile and determined the possible alternatives as A_1 Mercedes C 200d, A_2 Audi A4 2.0 TDI, A_3 Volkswagen Passat 2.0 TDI, A_4 Volvo S60 2.0 TDI, A_5 Opel Insignia 2.0 CDTI, A_6 BMW 320d, A_7 Ford Mondeo 2.0 TDCI, A_8 Toyota Avensis 2.0 D-3D and A_9 Peugeot 508 2.0.

There are a lot of automobile alternatives and conflicting criteria to be considered, so automobile selection is an important and difficult decision for the marble company. For selecting the best automobile for the general manager, a combined approach based on MACBETH and MULTI-MOORA methods are proposed. The weights of the decision criteria are determined with MACBETH method and then MULTI-MOORA method is used to determine the ranking of the alternatives.

First of all, decision criteria are defined and expressed in the form of a value tree as seen in Figure 1. These criteria are; C_1 Price (TL), C_2 Fuel consumption (lt/100km), C_3 Safety, C_4 Brand image, C_5 After sales service, C_6 Comfort, C_7 Design, C_8 Engine power (HP) and C_9 CO_2 Emissions (g/km).

Figure 1. MACBETH value tree



In order to determine the weights of the criteria with MACBETH method, criteria are entered into M-MACBETH software in descending order of their importance from left to right and top to bottom in the weighting matrix, as shown in Table 3.

In order to convert the performance levels for all criteria into proportionate quantitative MACBETH scores, they are pair-wise compared with the help of a seven point semantic scale. M-MACBETH software checked the consistency of these judgments and it is found that the entered judgments are consistent. Further, based on the provided differences of attractiveness, M-MACBETH software converts the ordinal performance levels into proportionate cardinal MACBETH scale using linear programming models. This MACBETH scale can be seen in the last column of the Table 3 and these values indicate the weights of the criteria. The weights of the criteria obtained with the MACBETH method can be seen in Table 4.

Table 3. Comparison of the criteria

Weighting (Automobile Selection)

	[C1]	[C2]	[C3]	[C4]	[C5]	[C6]	[C7]	[C8]	[C9]	[all lower]	Current scale
[C1]	no	very weak	weak	weak	moderate	strong	strong	v. strong	extreme	positive	20.00
[C2]		no	very weak	very weak	weak	moderate	moderate	strong	v. strong	positive	16.84
[C3]			no	very weak	weak	moderate	moderate	strong	v. strong	positive	15.79
[C4]				no	very weak	weak	moderate	strong	v. strong	positive	14.74
[C5]					no	very weak	weak	moderate	strong	positive	11.58
[C6]						no	very weak	weak	moderate	positive	8.42
[C7]							no	very weak	weak	positive	7.37
[C8]								no	very weak	positive	4.21
[C9]									no	positive	1.05
[all lower]										no	0.00

extreme
v. strong
strong
moderate
weak
very weak
no

Consistent judgements

Table 4. Weights of the criteria

Criteria	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9
Weights	0.20	0.1684	0.1579	0.1474	0.1158	0.0842	0.0737	0.0421	0.0105

After the weights of the criteria are determined by MACBETH method, MULTI-MOORA method is used to find the ranking of the automobile alternatives. According to MULTI-MOORA method, firstly decision matrix is formed as seen in Table 5. In this table, data of the automobile alternatives are given. The data for C_1, C_2, C_3, C_8 and C_9 are quantitative data whereas data for the C_4, C_5, C_6 and C_7 are qualitative data. The quantitative data like C_1 Price (TL), C_2 Fuel consumption (lt/100km), C_8 Engine power (HP) and C_9 CO_2 Emissions are obtained from the websites of the related automobiles [48, 49, 50, 51, 52, 53, 54, 55, and 56]. The data for C_3 that shows the safety values of the automobiles are obtained from Euro NCAP [57]. Euro NCAP has created the safety rating system to help consumers, compare automobiles more easily and to help them identify the safest choice for their needs. The safety rating is determined from a series of vehicle tests, designed and carried out by Euro NCAP. While obtaining the qualitative data, decision maker evaluated the alternatives by using 5 point scale in which 1=Poor, 2=Fair, 3=Good, 4=Very good, 5=Excellent. On the other hand, some of the automobile selection criteria have to be maximized and the others minimized. As seen in Table 5, C_3, C_4, C_5, C_6, C_7 and C_8 are maximization criteria and C_1, C_2, C_9 are minimization criteria. After forming the decision matrix, it is normalized by using Eq. (10) as shown in Table 6.

Table 5. Quantitative data for performance evaluation of alternatives

Optimization direction	min	min	max	max	max	max	max	max	min
Alternatives	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9
A_1	168000	4.2	35	5	5	5	5	136	109
A_2	179697	4.1	34.5	5	4	4	5	190	107
A_3	140600	4.5	33	3	3	3	3	150	119
A_4	134950	4.3	28	3	2	4	4	190	112
A_5	151980	5.6	35	3	4	4	3	170	147
A_6	181632	4	35	5	3	4	5	190	106
A_7	160620	4.8	33	2	3	3	2	180	125
A_8	162900	4.5	35.4	3	4	3	2	143	119
A_9	178000	4.2	32	2	3	2	3	180	110

Table 6. Normalized decision matrix

Alternatives	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9
A_1	0.344	0.312	0.348	0.458	0.470	0.456	0.445	0.265	0.309
A_2	0.368	0.304	0.343	0.458	0.376	0.365	0.445	0.370	0.303
A_3	0.288	0.334	0.328	0.275	0.282	0.274	0.267	0.292	0.337
A_4	0.276	0.319	0.279	0.275	0.188	0.365	0.356	0.370	0.317
A_5	0.311	0.416	0.348	0.275	0.376	0.365	0.267	0.331	0.416
A_6	0.372	0.297	0.348	0.458	0.282	0.365	0.445	0.370	0.300
A_7	0.329	0.356	0.328	0.183	0.282	0.274	0.178	0.351	0.354
A_8	0.333	0.334	0.352	0.275	0.376	0.274	0.178	0.279	0.337
A_9	0.364	0.312	0.318	0.183	0.282	0.183	0.267	0.351	0.311

Then, the weighted normalized decision matrix is formed by using Eq. (11) as shown in Table 7. From the weighted normalized decision matrix y_i^* values of each alternative are calculated with Eq. (12). These values can be seen in the last column of Table 7. And the ranking of alternatives with Ratio System of MOORA by considering y_i^* values are given in Table 8.

Table 7. Weighted normalized decision matrix and y_i^* values

Alternatives	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	y_i^*
A_1	0.069	0.053	0.055	0.068	0.054	0.038	0.033	0.011	0.003	0.135
A_2	0.074	0.051	0.054	0.068	0.044	0.031	0.033	0.016	0.003	0.116
A_3	0.058	0.056	0.052	0.041	0.033	0.023	0.020	0.012	0.004	0.063
A_4	0.055	0.054	0.044	0.041	0.022	0.031	0.026	0.016	0.003	0.067
A_5	0.062	0.070	0.055	0.041	0.044	0.031	0.020	0.014	0.004	0.067
A_6	0.074	0.050	0.055	0.068	0.033	0.031	0.033	0.016	0.003	0.107
A_7	0.066	0.060	0.052	0.027	0.033	0.023	0.013	0.015	0.004	0.033
A_8	0.067	0.056	0.056	0.041	0.044	0.023	0.013	0.012	0.004	0.061
A_9	0.073	0.053	0.050	0.027	0.033	0.015	0.020	0.015	0.003	0.031

Table 8. Ranking of alternatives with Ratio System of MOORA

Alternatives	Brands	y_i^*
A_1	Mercedes C 200d	0.135
A_2	Audi A4 2.0 TDI	0.116
A_6	BMW 320 d	0.107
A_5	Opel Insignia 2.0 CDTI	0.067
A_4	Volvo S60 2.0 TDI	0.067
A_3	Volkswagen Passat 2.0 TDI	0.063
A_8	Toyota Avensis 2.0 D-3D	0.061
A_7	Ford Mondeo 2.0 TDCI	0.033
A_9	Peugeot 508 2.0	0.031

In Reference Point Approach of MOORA, firstly reference points are determined for each criterion from the weighted normalized decision matrix. While determining reference points, highest values are choosen for maximization criteria (C_3, C_4, C_5, C_6, C_7 and C_8) and minimium values are choosen for the minimization criteria (C_1, C_2, C_9). These points are given at the last row of Table 9. Then, the distances between the alternatives and the reference points are calculated like in Table 9. Later, maximum values of these distances are determined and given in the last column of the Table 9. Then the best alternative is chosen by using the Tchebycheff Min-Max metric seen in Eq. (13). Final ranking of the alternatives are obtained by ranking the maximum distance values in increasing order like in Table 10.

Table 9. Reference Point Approach of MOORA

Alternatives	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	Max
A_1	0.014	0.003	0.001	0.000	0.000	0.000	0.000	0.004	0.000	0.014
A_2	0.018	0.001	0.001	0.000	0.011	0.008	0.000	0.000	0.000	0.018
A_3	0.002	0.006	0.004	0.027	0.022	0.015	0.013	0.003	0.000	0.027
A_4	0.000	0.004	0.012	0.027	0.033	0.008	0.007	0.000	0.000	0.033
A_5	0.007	0.020	0.001	0.027	0.011	0.008	0.013	0.002	0.001	0.027
A_6	0.019	0.000	0.001	0.000	0.022	0.008	0.000	0.000	0.000	0.022
A_7	0.011	0.010	0.004	0.041	0.022	0.015	0.020	0.001	0.001	0.041
A_8	0.011	0.006	0.000	0.027	0.011	0.015	0.020	0.004	0.000	0.027
A_9	0.018	0.003	0.005	0.041	0.022	0.023	0.013	0.001	0.000	0.041
Reference points	0.055	0.050	0.056	0.068	0.054	0.038	0.033	0.016	0.003	

Table 10. Ranking of alternatives with Reference Point Approach of MOORA

Alternatives	Brands	Max
A_1	Mercedes C 200d	0.014
A_2	Audi A4 2.0 TDI	0.018
A_6	BMW 320 d	0.022
A_3	Volkswagen Passat 2.0 TDI	0.027
A_5	Opel Insignia 2.0 CDTI	0.027
A_8	Toyota Avensis 2.0 D-3D	0.027
A_4	Volvo S60 2.0 TDI	0.033
A_7	Ford Mondeo 2.0 TDCI	0.041
A_9	Peugeot 508 2.0	0.041

In Full Multiplicative Form of MOORA, multiplicative ranking index U_i for each alternative are determined with Eq. (14) and these values are given at last column of Table 11. Then ranking of the alternatives are obtained by ranking U_i values in descending order. This ranking obtained by Full Multiplicative Form of MOORA can be seen in Table 12.

Table 11. Full Multiplicative Form of MOORA

Alternatives	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	U_i
A_1	11.093	1.273	1.753	1.268	1.205	1.145	1.126	1.230	1.050	0.286
A_2	11.244	1.268	1.749	1.268	1.174	1.124	1.126	1.247	1.050	0.274
A_3	10.705	1.288	1.737	1.176	1.136	1.097	1.084	1.235	1.051	0.235
A_4	10.618	1.278	1.692	1.176	1.084	1.124	1.108	1.247	1.051	0.235
A_5	10.873	1.337	1.753	1.176	1.174	1.124	1.084	1.241	1.054	0.239
A_6	11.268	1.263	1.753	1.268	1.136	1.124	1.126	1.247	1.050	0.267
A_7	10.994	1.302	1.737	1.108	1.136	1.097	1.052	1.244	1.052	0.208
A_8	11.025	1.288	1.756	1.176	1.174	1.097	1.052	1.232	1.051	0.231
A_9	11.222	1.273	1.728	1.108	1.136	1.060	1.084	1.244	1.051	0.207

Table 12. Ranking of alternatives with Full Multiplicative Form of MOORA

Alternatives	Brands	U_i
A_1	Mercedes C 200d	0.286
A_2	Audi A4 2.0 TDI	0.274
A_6	BMW 320 d	0.267
A_5	Opel Insignia 2.0 CDTI	0.239
A_3	Volkswagen Passat 2.0 TDI	0.235
A_4	Volvo S60 2.0 TDI	0.235
A_8	Toyota Avensis 2.0 D-3D	0.231
A_7	Ford Mondeo 2.0 TDCI	0.208
A_9	Peugeot 508 2.0	0.207

The final ranking of the alternatives are determined with MULTI-MOORA method. By using dominance theory one final ranking is obtained from three ranks of Ratio System, Reference Point Approach and Full Multiplicative Form. This final ranking is given in Table 12. According to the result of MULTI-MOORA, the final ranking of alternatives is obtained as $A_1>A_2>A_6>A_5>A_3>A_4>A_8>A_7>A_9$ and the best automobile alternative is A1 Mercedes C 200d

Table 13. Ranking of alternatives with MULTI- MOORA

Ratio System of MOORA	Reference Point Approach of MOORA	Full Multiplicative Form of MOORA	MULTI-MOORA
A_1	A_1	A_1	A_1
A_2	A_2	A_2	A_2
A_6	A_6	A_6	A_6
A_5	A_3	A_5	A_5
A_4	A_5	A_3	A_3
A_3	A_8	A_4	A_4
A_8	A_4	A_8	A_8
A_7	A_7	A_7	A_7
A_9	A_9	A_9	A_9

5. CONCLUSION

In this paper, a combined approach based on MACBETH and MULTI-MOORA methods is proposed for the first time and its applicability is illustrated with an automobile selection problem of a marble company. Decision criteria are determined as price, fuel consumption, safety, brand image, after sales service, comfort, design, engine power and CO_2 emissions. The weights of these criteria are determined with MACBETH method. Then the final ranking of the automobile alternatives is obtained with MULTI-MOORA method as $A_1>A_2>A_6>A_5>A_3>A_4>A_8>A_7>A_9$. As the best alternative is A_1, it is advised to the marble company to purchase Mercedes C 200d for the general manager. They found the results satisfactory and decided to purchase Mercedes C 200d.

In this paper, combined approach is based on MACBETH and MULTI-MOORA methods. These two MCDM methods are preferred because of their advantages over others. MACBETH has advantage because it only requires qualitative judgements to weight the criteria and to score the alternatives. Also, the support of M-MACBETH software improves the usefulness of this method in solving complex decision-making problems [20]. On the other hand, this software provides a consistency checking and if the judgements are found inconsistent, it suggests possible alterations to make them consistent.

MULTI-MOORA method combines the results of three MOORA approach namely Ratio System, Reference Point Method and Full Multiplicative Form. For this reason it gives guarantee for robustness [41]. This is the advantage of MULTI-MOORA over other MCDM methods. Also, Full Multiplicative Form of MOORA does not need the use of normalization. This reduces the amount of calculations required and saves time.

In future studies, proposed combined approach can be used to solve different MCDM problems of the companies. And the weights of the criteria can be determined with AHP (Analytic Hierarch Process) method instead of MACBETH. Also other MCDM methods can be used for ranking the alternatives and the obtained results can be compared.

References

[1]. Güngör, İ., İşler D. B. (2004). Analitik Hiyerarşi Yaklaşimi İle Otomobil Seçimi, *ZKÜ Sosyal Bilimler Dergisi*, 1(2), 21-33.

[2]. Terzi, Ü., Hacaloğlu, S.E., Aladağ, Z. (2006). Otomobil Satın Alama Problem için Bir Karar Destek Modeli, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi,* 5(10), 43-49.

[3]. Soba, M. (2012). PROMETHEE Yöntemi Kullanarak En Uygun Panelvan Otomobil Seçimi ve Bir Uygulama, *Journal of Yasar University*, 28(7), 4708 – 4721.

[4]. Sakthivel, G. Ilangkumaran, M. Nagarajan, G. Raja, A. Ragunadhan P.M., Prakash, J. (2013.) A Hybrid MCDM Approach for Evaluating an Automobile Purchase Model, *International Journal of Information and Decision Sciences*, 5(1), 50-85.

[5]. Yavaş, M., Ersöz, T., Kabak, M. (2014). Otomobil Seçimine Çok Kriterli Yaklaşım Önerisi, *İşletme ve İktisat Çalışmaları Dergisi*, 2(4), 110-118.

[6]. Burgazoğlu, H, (2015). MACBETH, B.F. Yıldırım, E. Önder (Ed.), *Çok Kriterli Karar Verme Yöntemleri*, Bursa: Dora Yayınları.

[7]. Bana e Costa C.A., Chagas M. P. (2004). A Career Choice Problem: An Example of How to Use MACBETH to Build a Quantitative Value Model Based on Qualitative Value Judgments, *European Journal of Operational Research*, 153, 323–331.

[8]. Karande P., Chakraborty, S. (2014). A Facility Layout Selection Model Using MACBETH Method, *Proceedings of the 2014 International Conference on Industrial Engineering and Operations Management*, January 7 – 9, Bali, Indonesia.

[9]. Bana e Costa C.A., Ensslin, L., Correa E.C., Vansnick, J.C. (1999). Decision Support Systems in Action: Integrated Application in a Multicriteria Decision Aid Process, *European Journal of Operational Research*, 113, 315-335.

[10]. Bana e Costa C.A, Nunes da Silva, F.,Vansnick, J.C. (2001). Conflict Dissolution in the Public Sector: A case-study, *European Journal of Operational Research*, 130, 388-401.

[11]. Bana e Costa, C.A. (2001). The Use of Multi-Criteria Decision Analysis to Support the Search for Less Conflicting Policy Options in a Multi-Actor Context: Case Study, *Journal of Multi-Criteria Decision Analysis,* 10, 111–125.

[12]. Bana e Costa C.A., Oliveira R.C. (2002). Assigning Priorities for Maintenance, Repair and Refurbishment in Managing a Municipal Housing Stock, *European Journal of Operational Research*, 138, 380–391.

[13]. Bana e Costa C.A, Costa-Lobo M.L., Ramos, I.A., Vansnic, J.C. (2002a). Multicriteria Approach for Strategic Town Planning. In D. Bouyssou et al. (Ed.), *Aiding Decisions with Multiple Criteria*, Kluwer Academic Publishers.

[14]. Bana e Costa, C.A., Barroso, L.A., Soares, J.O. (2002b). Qualitative Modelling of Credit Scoring: A Case Study in Banking, *European Research Studies,* 5 (1-2), 37–51.

[15]. Bana e Costa C.A., Correa, E.C., De Corte, J.-M.,Vansnic , J.C. (2002c). Facilitating Bid Evaluation in Public Call for Tenders: A Socio-Technical Approach, *Omega*, 30, 227 – 242.

[16]. Roubens, M., Rusinowska, A., De Swart, H. (2006). Using MACBETH to Determine Utilities of Governments to Parties in Coalition Formation, *European Journal of Operational Research*, 172, 588–603.

[17]. Cliville V., Berrah L., Mauris, G. (2007). Quantitative Expression and Aggregation of Performance Measurements Based on the MACBETH Multi-Criteria Method, *International Journal of Production Economics*, 105, 171–189.

[18]. Montignac, F., Noirot, I., Chaudourne, S. (2009). Multi-Criteria Evaluation of On-Board Hydrogen Storage Technologies Using the MACBETH Approach, *International Journal hydrogen energy*, 34, 4561 – 4568.

[19]. Fakhfakh, N., Verjus, H., Pourraz, F., Moreaux, P. (2011). Measuring the Satisfaction Degree of Quality Attributes Requirements for Services Orchestrations, *Proceedings of 4th International Conference on Communication Theory, Reliability, and Quality of Service*, Hungary, pp. 89-94.

[20]. Karande, P., Chakraborty, S. (2013). Using MACBETH Method for Supplier Selection in Manufacturing Environment, *International Journal of Industrial Engineering Computations*, 4, 259–272.

[21]. Rodrigues, T.C. (2014). The MACBETH Approach to Health Value Measurement: Building a Population Health Index in Group Processes, *Procedia Technology*, 16, 1361 – 1366.

[22]. Brauers, W.K.M., Zavadskas, E.K. (2006). The MOORA Method and Its Application to Privatization in a Transition Economy. *Control and Cybernetics,* 35(2), 445-469.

[23]. Brauers, W.K.M., Zavadskas, E.K., Turskis, Z. Vilutiene, T. (2008). Multi-Objective Contractor's Ranking by Applying the MOORA Method. *Journal of Business Economics and Management,* 9(4), 245-255.

[24]. Brauers, W.K.M., Zavadskas, E.K., Peldschus, F., Turskis, Z. (2008). Multi-Objective Decision Making for Road Design, *Transport,* 23(3), 183-193.

[25]. Kalibatas, D., Turskis Z. (2008). Multicriteria Evaluation of Inner Climate by Using MOORA Method. *Information Technology and Control*, 37(1), 79-83.

[26]. Brauers, W.K.M., Ginevičius, R. (2009). Robustness in Regional Development Studies: The Case of Lithuania. *Journal of Business Economics and Management, 10(2)*, 121–140.

[27]. Brauers, W.K.M., Ginevičius, R., Podvezko, V. (2010). Regional Development in Lithuania Considering Multiple Objectives by The MOORA Method. *Technological and Economic Development of Economy*, 16(4), 613-640.

[28]. Brauers, W.K.M., Zavadskas, E.K. (2010). Project Management by Multimoora as an Instrument for Transition Economies. *Technological and Economic Development of Economy*, 16(1), 5-24.

[29]. Gadakh, V. S. (2011). Application of MOORA Method for Parametric Optimization of Milling Process. *International Journal of Applied Engineering Research*, 1 (4), 743-758.

[30]. Archana, M., Sujatha, V. (2012). Application of Fuzzy MOORA and GRA in Multi Criterion Decision Making Problems, *International Journal of Computer Applications*, 53(9), 46-50.

[31]. Karaca, T. (2011). *Proje Yönetiminde Çok Kriterli Karar Verme Teknikleri Kullanarak Kritik Yolun Bulunması*. (Yayımlanmamış yüksek lisans tezi). Gazi Üniversitesi/ Fen Bilimleri Enstitüsü, Ankara.

[32]. Önay, O., Çetin, E. (2012). Turistik Yerlerin Popülaritesinin Belirlenmesi: İstanbul Örneği. *Yönetim: İstanbul Üniversitesi İşletme İktisadı Enstitüsü Dergisi*, 23(72), 90-109.

[33]. Dey, B., Bairagi, B., Sarkar, B., Sanyal, S. (2012). A Moora Based Fuzzy Multi Criteria Decision Making Approach for Supply Chain Strategy Selection. *International Journal of Industrial Engineering Computations*, *3*, 649-662.

[34]. Mandal, U. K.ve Sarkar, B. (2012). Selection of Best Intelligent Manufacturing System Under Fuzzy MOORA Conflicting MCDM Environment. *International Journal of Engineering Technology and Advanced Engineering*, 2(9), 301-310.

[35]. Balezantis, A., Balezantis, T., Brauers, W.K.M. (2012). Multimoora-FG: A Multi-Objective Decision Making Method for Linguistic Reasoning with an Application to Personnel Selection. *Informatica*, 23(2), 173-190.

[36]. Tepe, S., Görener, A. (2014). Analitik Hiyerarşi Süreci ve MOORA Yöntemlerinin Personel Seçiminde Uygulanması. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi, 13(25)*, 1-14

[37]. Görener, A., Dinçer, H., Hacıoğlu, Ü. (2013). Application of Multi-Objective Optimization on the Basis of Ratio Analysis (MOORA) Method for Bank Branch Location Selection. *International Journal of Finance & Banking Studies*, 2(2), 41-52.

[38]. Yıldırım, B.F., Önay, O. (2013). Bulut Teknolojisi Firmalarının Bulanık AHP- MOORA Yöntemi Kullanılarak Sıralanması, *İ.Ü.İşletme Fakültesi, İşletme İktisadı Enstitüsü Yönetim Dergisi,* 24(75), 59-81.

[39]. Obayiuwana, E. Falowo, O. (2015) A Multimoora Approach to Access Network Selection Process in Heterogeneous Wireless Networks, AFRICON 2015, 14-17 September, Addis Ababa, IEEE, pp. 1-5, doi.10.1109/AFRCON. 2015.7331973

[40]. Brauers W. K. M., Zavadskas E.K. (2011). Multimoora Optimization Used to Decide on a Bank Loan to Buy Property, *Technological and Economic Development of Economy*, 17(1), 174-188.

[41]. Brauers, W.K.M., Ginevičius R. (2010). The Economy of The Belgian Regions Tested with Multimoora, *Journal of Business Economics and Management*, 11:2, 173-209.

[42]. Brauers, W. K. M., Baležentis A., Baležentis, T. (2011).Multimoora for the EU Member States Updated with Fuzzy Number Theory, *Technological and Economic Development of Economy*, 17:2, 259-290.

[43]. Baležentis A., Baležentis, T., Brauers W. K. M., (2012). MULTIMOORA-FG: A Multi-Objective Decision Making Method for Linguistic Reasoning with an Application to Personnel Selection, *Informatica*, 23(2), 173–190.

[44]. Brauers W. K. M., Zavadskas E.K. (2012). Robustness of MULTIMOORA: A Method for Multi-Objective Optimization, *Informatica,* 23(1), 1–25.

[45]. Datta S., Sahu, N., Mahapatra, S. (2013). Robot Selection Based on Grey-MULTIMOORA Approach, *Grey Systems: Theory and Application*, 3(2), 201 – 232.

[46]. Baležentis T., Baležentis, A. (2014). A Survey on Development and Applications of the Multi-criteria Decision Making Method MULTIMOORA, *Journal Of Multi-Criteria Decision Analysis*, 21, 209–222.

[47]. Aksoy, E. (2015). AHP Temelli MULTIMOORA ve COPRAS Yöntemi ile Türkiye Kömür İşletmeleri'nin Performans Değerlendirmesi , *Hacettepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 33(4), 1-28.

[48]. http://www.mercedes-benz.com.tr/content/turkey/mpc/mpc_turkey_website/tr/home_mpc/passengercars/home/new_cars/models/c-class/c205/advice_sales/brosur.html [17.01.2016]

[49]. http://www.audi.com.tr/tr/brand/tr/models/a4/a4_sedan/teknik_veriler/boyutlar.html#source=http://www.audi.com.tr/tr/brand/tr/models/a4/a4_sedan/teknik_veriler/teknik_detay.html &container=page [17.01.2016]

[50]. http://binekarac.vw.com.tr/pdf/passat.pdf [17.01.2016]

[51]. http://www.volvocars.com/tr/modeller/new-models/s60/ozellikler [17.01.2016]

[52]. http://www.opel.com.tr/content/dam/Opel/Europe/turkey/nscwebsite/tr/11_Brochure_Library/02_Specsheets/Insignia_MY16.5_FINAL.pdf [17.01.2016]

[53]. http://www.bmw.com.tr/content/dam/bmw/marketTR/bmw_com_tr/pdf/bmw_3_serisi_sedan_katalog_f30.pdf.asset.1457705813418.pdf [17.01.2016]

[54]. https://apps.ford.com.tr/ebrochure/brochurefiles/mondeo.pdf [17.01.2016]

[55]. https://www.toyota.com.tr/new-cars/avensis/index.json#/publish/compare_engines/selection=2WW-20D1-1436M [17.01.2016]

[56]. http://www.peugeot.com.tr/media/deliacms/media//74/7412-88db21.pdf [17.01.2016]

[57]. http://www.euroncap.com/tr [18.01.2016]

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2016.04.01.OR.03

PRODUCTIVITY CHANGE: AN EMPIRICAL STUDY ON TURKISH STATE UNIVERSITIES

F. Cengiz DİKMEN *

Faculty of Economics and Business Administration, Kocaeli University, Kocaeli, Turkey

Received: 11 January 2016
Accepted: 12 May 2016

Abstract

As it is known universities are public institutions providing educational and training services. They are also engaged with research activities. The services provided by these institutions concerns very closely both the public opinion and the public officials from numerous aspects. Thus, the resources allocated to the activities of these institutions must be evaluated to what extent it is used efficiently. In addition, the development of the institutions over time is also noteworthy. In this context, a DEA-based approach known as MPI (Malmquist Productivity Index) is used to evaluate the efficiency of state universities and to reveal the technological change and "catching-up" over time if there exists. MPI is a method of measuring the influence of time shift. It is designed to calculate the efficient frontier shift in a certain period of time. The efficiency shifts between two periods of time give the institutions the opportunity to compare and evaluate their relative competitive positions. This study comprises two academic periods, namely, 2000/01 and 2009/10 in order to investigate the productivity change on a sample from the state universities of Turkey.

Keywords: Total factor productivity, Malmquist Productivity Index, Data Envelopment Analysis, Turkish State universities

VERİMLİLİK DEĞİŞİMİ: TÜRK DEVLET ÜNİVERSİTELERİ ÜZERİNE AMPİRİK BİR ÇALIŞMA

Özet

Bilindiği üzere üniversiteler eğitim ve öğretim hizmeti veren kamu kuruluşlarıdır. Üniversiteler ayrıca araştırma faaliyetleriyle de ilgilidirler. Bu kurumlar tarafından sağlanan hizmetler çeşitli açılardan kamuoyu ve resmi kurumları çok yakından ilgilendirmektedir. Dolayısıyla, bu kurumların faaliyetlerine tahsis edilen kaynakların ne ölçüde verimli kullanıldığının değerlendirilmesi gerekmektedir. Bunun yanında, bu kurumların zaman içindeki gelişimi de dikkat çekici bir konudur. Bu bağlamda bu çalışmada, MVE (Malmquist Verimlilik Endeksi) olarak bilinen DEA-temelli bir yaklaşım, devlet üniversitelerinin etkinliğini değerlendirmek ve varsa zaman içindeki teknolojik değişimini ve etkinlik sınırına yakınlığını ortaya çıkarmak için kullanılmaktadır. MVE zaman değişiminin etkisini ölçmede kullanılan bir yöntemdir. MVE belli bir zaman diliminde etkin üretim sınırındaki kaymayı hesaplamak için geliştirilmiştir. İki zaman dilimi arasındaki etkinliğin ölçümü kurumlara göreli yarışmacı konumlarını değerlendirme ve karşılaştırma fırsatı tanımaktadır. Bu çalışma, Türk devlet üniversitelerinden oluşan bir örneklem üzerinden 2000/01 ve 2009/10 akademik dönemlerindeki verimlilik değişimini incelemeyi amaçlamaktadır.

Anahtar Kelimeler : Toplam faktör verimliliği, Malmquist Verimlilik Endeksi, Veri Zarflama Analizi, Türk Devlet üniversiteleri

* fcdikmen@kocaeli.edu.tr (Corresponding author)

1. INTRODUCTION

As it is known universities are institutions engaged in education and training, besides the research activities. The services provided by these institutions are directly related to public officials and the public opinion in many ways. Turkey has a growing young population demanding high level education every year. While the governments allocate big budgets for these public services, governments have the responsibility to be transparent and accountable to budgetary expenditures. Being transparent and accountable to the public necessitates controlling and pursuing efficiency and productivity in the allocation and management of public resources [2]. As a public institution the universities have also the responsibility to take necessary steps to be efficient while expending the public budget and to pursue productivity every year. Besides, the development of the institutions over time is also noteworthy. Higher education is one of the main sources of economic growth. Thus inefficiency in the university sector may cause a real welfare loss as does the misallocation of resources elsewhere in the economy. So it is vital to design and make improvements in educational policy which may lead to higher economic growth [1].

Efficiency can be measured either for a specific point of time or to evaluate the change of efficiency between two or more time periods. The latter points regress or progress in the efficiency if exists. In this study the purpose is to track the productivity change in Turkish public universities by taking into account the changes both in efficiency and technology.

In measuring productivity change for a specific set of decision making unit (DMU) between two or more time periods, two approaches have been used: (i) the econometric estimation of a production function, (ii) the construction of index numbers. While the former approach necessitates a functional form on the structure of production technology the latter approach does not require it. Once one decides to use the index number approach he/she may have three choices: (i) Fisher index, (ii) Törnqvist index, (iii) Malmquist index [15]. In this study Malmquist index approach is selected in measuring the productivity change in higher education. This selection highly depends on two reasons. Malmquist index rests exclusively on quantity information which means that it does not require price information. Because, obtaining price information for high level education institutions would be inappropriate regarding the outputs and inputs of these organizations. Particularly, in the analysis of the growth for public sector institutions, the data about the prices of inputs and outputs does not exist. For example, if one considers the inputs (academic or non-academic staff), and the outputs (number of published articles, graduate or post-graduate students) of higher education institutions, what price can be assigned to them? Secondly, it is a non-parametric approach which does not require a functional form for the production technology. Besides, the Malmquist approach gives the opportunity to decompose the productivity change into constituent sources of this change as technical and efficiency changes [15]. Thus, Malmquist index of total factor productivity change has gained great popularity in recent years.

In the following sections, Malmquist total productivity index and data envelopment methodology will be introduced and in the last section the Malmquist approach will be applied to a sample of Turkish public universities to measure productivity change with the aid of spreadsheet.

2. MALMQUIST INDEX

A total factor productivity index measures the change in total output relative to the change in the usage of all inputs. The change, if exists can be decomposed into two components, namely, the change in technical efficiency, and the change in technology. While the former change shows the relative closeness to the efficiency frontier, the latter change shows the shift in the efficiency frontier. It is an index representing Total Factor Productivity (TFP) growth of a DMU, in that it reflects progress or regress in efficiency along with progress or regress of the frontier technology over time under the multiple inputs and multiple outputs framework. Total factor productivity index relies on the works of Malmquist [18], Caves et al. [5], Fare et al. [11], Fare et al. [12] chronologically [4].

The Malmquist index is calculated by using distance functions. First a production frontier is constructed using data on multiple inputs and multiple outputs of all the DMUs in the sample in time periods *t* and *t+1*. In the next step, the radial distance for a specific DMU is computed relative to the production frontier in time *t* and in time *t+1* constructed in the first step. For the MPI to be computed we need to calculate four distances. Two distance measures for two single periods and other two distance measures for two mixed periods. Supposing that we have *n* DMUs, with each using *m* inputs, $x_i = 1,\dots,m$ and produce *s* outputs $y_r = 1,\dots,s$. If $d_0^t\left(x_0^t, y_0^t\right)$ is the distance measure for a specific DMU_0 for time period *t*, $d_0^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)$ for time period *t+1* and if $d_0^t\left(x_0^{t+1}, y_0^{t+1}\right)$ is the distance from period *t+1* observation to period *t* frontier and $d_0^{t+1}\left(x_0^t, y_0^t\right)$ is the distance from period *t* observation to period *t+1* frontier then input-oriented Malmquist productivity index can be expressed as:

$$= M_0\left(x_t, y_t, x_{t+1}, y_{t+1}\right) = \left[\frac{d_0^t\left(x_0^{t+1}, y_0^{t+1}\right)}{d_0^t\left(x_0^t, y_0^t\right)} x \frac{d_0^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)}{d_0^{t+1}\left(x_0^t, y_0^t\right)}\right]^{0.5} \quad (1)$$

M_0 measures the productivity change between periods *t* and *t+1*. If M_0 takes the value one, it indicates that there is no change (stagnation) in productivity. In case the value is greater than unity, it indicates an improvement (growth) in productivity and if it is less than unity it indicates declination in productivity. The distances measured may be either input or output oriented, and accordingly the Malmquist indices may give different results. But if it is assumed that the underlying production technology exhibits constant returns to scale (CRS) for the time periods, then the input and output oriented Malmquist TPF indices are equal [8; 21].

Equation (2) can be equivalently expressed as:

$$= M_0\left(x_t, y_t, x_{t+1}, y_{t+1}\right) = \frac{d_0^{t+1}\left(x_{t+1}, y_{t+1}\right)}{d_0^t\left(x_t, y_t\right)} x \left[\frac{d_0^t\left(x_0^{t+1}, y_0^{t+1}\right)}{d_0^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)} x \frac{d_0^t\left(x_0^t, y_0^t\right)}{d_0^{t+1}\left(x_0^t, y_0^t\right)}\right]^{0.5}$$

(2)

The equation shows the decomposition of the Malmquist index into a product of two measures: (I) the change in the technical efficiency and (ii) the geometric mean of the change in the frontier. The first part of the index measures the change in technical efficiency between two time periods [20]. This component is also termed as "catch-up" which compares the closeness of DMU_0 in each period's efficient boundary. For values greater than one indicates an improvement in relative technical efficiency during the period considered. The ratio inside the bracket measures boundary shift. A value greater than unity will indicate that there is a technological progress in the industry the DMUs operate. Precisely, the decomposition of the index as in equation (2), provides valuable information about the sources of the overall productivity change.

In order to capture the impact of any scale size changes on productivity, equation (2) can further be decomposed into pure technical efficiency change, scale efficiency change and technological change components which was first put forth by Fare et al.(1994b) and can be stated as in equation (3) [21]**.** The first component, the pure technical efficiency change, is the ratio of the efficiency measured at time *t+1* to the efficiency measured at time *t* under variable returns to scale (VRS) assumption, which in turn is termed as pure technical efficiency catch-up. It is interpreted as the technical efficiency term in equation (2). The only difference is the assumption of the production technology used as either CRS or VRS. The second component, scale efficiency change, indicates to what extent the DMU_0 has become more scale efficient (SCE) between two time periods. Therefore it captures the impact of any change in scale size of DMU_0 on its productivity [21]. Scale efficiency is the ratio of two efficiency scores which are measured under two production technologies, namely, VRS and CRS at a point of time. Thus, the scale efficiency catch-up is the proportion of these two scale efficiencies measured at two different point of time. Consequently, M_0 can be expressed as:

$$= M_0\left(x_t, y_t, x_{t+1}, y_{t+1}\right) = \frac{d_{0-VRS}^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)}{d_{0-SCE}^t\left(x_0^t, y_0^t\right)} \cdot \frac{d_{0-SCE}^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)}{d_{0-SCE}^t\left(x_0^t, y_0^t\right)} \cdot \left[\frac{d_0^t\left(x_0^{t+1}, y_0^{t+1}\right)}{d_0^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right)} \cdot \frac{d_0^t\left(x_0^t, y_0^t\right)}{d_0^{t+1}\left(x_0^t, y_0^t\right)}\right]^{0.5} \quad (3)$$

This component can attain a value greater than, equal to, or less than unity according to the DMU0's scale size contribution to productivity change. A value of greater than unity indicates that scale size has positive impact on the productivity change, which means that the DMU0 is in the direction of technical optimal scale. On the contrary if it is less than unity scale size has negative impact on productivity change meaning that DMU0 is in the direction away from the technical optimal scale [15].

3. DEA and MALMQUIST INDEX

DEA was originated by the seminal paper of Charnes, Cooper and Rhodes in [6]. Since then DEA has become a well-known technique to deal with efficiency and productivity measurement. It was originally designed to measure of a set of homogeneous decision making units like universities, hospitals and schools which are non-profit organizations. However, later on, more and more DEA research has been adopted and applied to measure the performance of profit organizations.

It is a non-parametric technique in the sense that it does not require a priori specification of input and output weights. DEA applications generally use cross-section data to measure performance of DMUs that is the performance at the same time of point. DEA can also calculate the productivity change of a DMU over time. Thus it became applicable to panel data to measure the productivity changes between two time periods [7].

Linear programming is the underlying methodology that makes DEA particularly powerful compared with alternative productivity management tools. Fare et al. [1992] have used DEA model as a mathematical programing-based methodology to compute Malmquist index of productivity change which is applied to Swedish pharmacies. After this seminal study, there have been a considerable number of studies in the literature about the framework, decomposition and computation of the Malmquist index using DEA approach. DEA-based Malmquist productivity index has been used extensively in diverse scientific and economic fields. One can mention the relevant studies in literature as: Fare et al. [10] investigate the productivity changes in Swedish pharmacies between 1980-1989; Fare et al [10] employ MPI to investigate the productivity change in Swedish hospitals; Fare et. [12] analyzes productivity growth in 17 OECD countries over the period 1979-1988; Şentürk [19] uses MPI to estimate and analyze the TFP growth rates of public and private Turkish manufacturing industries over the period 1985 to 2001 using DEA linear programming technique; Flegg et al. [13] uses MPI to examine the technical efficiency of 45 British universities in the period 1980/81-1992/93; Kao Chiang and Liu Shinang-Tai [16] measures the efficiency of 22 Taiwanese commercial banks for the period 2009-2011; Brennan et al.[3] applies the methodology to analyze productivity of Dutch schools using 2002-2007 data; Forsund Finn R. and Edvordson [14] studies the performance of local taxes overtime using DEA to calculate MPI; Yi-Hsing et al. [17] investigates relative efficiency of management and variation of managerial efficiency among 37 domestic banks in Taiwan and so forth.

Following Fare et al. [10] in order to compute Malmquist index four efficiency calculations are needed. Two for two single time periods *t* and *t+1* and two for two mixed periods. The two sing le period measures can be obtained by the DEA_{CRS} model given below. Efficiency models for time *t* and *t+1* are respectively can be stated as:

$$\begin{aligned} &\theta_0^t\left(x_0^t, y_0^t\right) = \min \theta_0 \\ &s.t. \\ &\sum_{j=1}^{n} \lambda_j x_j^t \le \theta_0 x_0^t \qquad \text{(M.1)} \\ &\sum_{j=1}^{n} \lambda_j y_j^t \ge y_0^t \\ &\lambda_j \ge 0,\ j = 1,\dots, n \end{aligned}$$

$$\begin{aligned} &\theta_0^{t+1}\left(x_0^{t+1}, y_0^{t+1}\right) = \min \theta_0 \\ &s.t. \\ &\sum_{j=1}^{n} \lambda_j x_j^{t+1} \le \theta_0 x_0^{t+1} \qquad \text{(M.2)} \\ &\sum_{j=1}^{n} \lambda_j y_j^{t+1} \ge y_0^{t+1} \\ &\lambda_j \ge 0,\ j = 1,\dots, n \end{aligned}$$

And the efficiency models for DMU_o for the two mixed periods; the first model of the two, compares x_0^{t+1} data to the production technology (boundary) at time *t* and the second model compares x_0^t data to boundary at time *t+1* can be stated as:

$$\begin{aligned} &\theta_0^t\left(x_0^{t+1}, y_0^{t+1}\right) = \min \theta_0 \\ &s.t. \\ &\sum_{j=1}^{n} \lambda_j x_j^t \le \theta_0 x_0^{t+1} \qquad \text{(M.3)} \\ &\sum_{j=1}^{n} \lambda_j y_j^t \ge y_0^{t+1} \\ &\lambda_j \ge 0,\ j = 1,\dots, n \end{aligned}$$

$$\begin{aligned} &\theta_0^{t+1}\left(x_0^t, y_0^t\right) = \min \theta_0 \\ &s.t. \\ &\sum_{j=1}^{n} \lambda_j x_j^{t+1} \le \theta_0 x_0^t \qquad \text{(M.4)} \\ &\sum_{j=1}^{n} \lambda_j y_j^{t+1} \ge y_0^t \\ &\lambda_j \ge 0,\ j = 1,\dots, n \end{aligned}$$

4. PRODUCTIVITY CHANGE IN TURKISH UNIVERSITIES

The data consists of annual observation of a sample of 37 Turkish public universities for the period between 2000/01 and 2009/10. There are total of 185 universities. 109 of them are the public and 76 of them are the private university. Private universities are excluded out of the scope of the analysis due to the lack of appropriate data for the mentioned time periods.

Three categories of output are used in the analysis: (i) undergraduate completions, (ii) postgraduate completions and (ii) published articles. The inputs included in the analysis are full time equivalent (i) academic staff, (ii) non-academic staff. The input and output specifications are consistent with the studies in the literature.

The input and output data used in the analysis were obtained from the Council of Higher Education of Turkey for the time period 2009/10. For the time period 2000/01 the input and output data is obtained by scanning the annual reports of the universities sent to the Council of Higher Education.

In the current study models M.1, M.2, M.3 and M.4 are employed to evaluate the productivity changes in high education sector. Models M.1 and M.2 allows to determine the technical efficiency (TE) of each university for each academic year assuming CRS technology. The results are depicted in Table 1 in the third and fourth columns under the heading TE_t and TE_{t+1}. The fifth and sixth columns denote the technical efficiency results assuming VRS technology. The last two columns are the technical efficiency results for each university for the mixed time periods. Considering the arithmetic mean and the standard deviation for the two time periods we can express that the efficiency scores of the universities tend to rise from the first period to the second period of time. Also the decline in standard deviation indicates a reduction in the amount of variation in performance across the university sector.

TE scores for the two time period plotted in Figure 1 supports the upward trend in efficiency scores.

Table 1. Efficiency scores for the universities between 2000/01 and 2009/10

DMU	University	TE_t	TE_{t+1}	Te_t(VRS)	TE_{t+1}(VRS)	$TE_{t\text{-üretim}}$	$Te_{t+1\text{-üretim}}$
1	ADNAN MENDERES	0,296	0,655	0,462	0,676	1,226	1,449
2	ATATÜRK	0,462	0,622	0,856	0,649	1,996	0,997
3	BALIKESİR	0,646	1,000	0,826	1,000	1,913	0,624
4	CELAL BAYAR	1,000	0,747	1,000	0,761	2,191	6,393
5	CUMHURİYET	0,483	0,539	0,598	0,553	1,281	1,797
6	ÇUKUROVA	0,540	0,715	0,682	0,749	3,511	1,987
7	DİCLE	0,285	0,529	0,295	0,584	2,163	0,905
8	DOKUZ EYLÜL	0,404	0,479	0,575	0,536	2,014	0,923
9	DUMLUPINAR	1,000	1,000	1,000	1,000	5,951	1,499
10	ERCİYES	0,434	1,000	0,557	1,000	2,596	1,255
11	BOĞAZİÇİ	0,846	0,766	0,896	0,840	3,417	2,500
12	GALATASARAY	0,263	0,987	1,000	1,000	5,796	1,023
13	GAZİANTEP	0,657	0,816	0,718	0,989	1,503	2,448
14	GAZİOSMANPAŞA	1,000	0,754	1,000	0,853	1,455	6,081
15	HACETTEPE.	0,390	0,683	0,794	1,000	1,485	1,485
16	HARRAN	0,450	0,597	0,615	0,741	1,545	1,114
17	İSTANBUL	0,595	0,827	1,000	1,000	4,239	0,527
18	KARADENİZ TEKNİK	0,570	0,686	0,916	0,728	1,686	1,145
19	ODTÜ	0,755	1,000	0,758	1,000	4,154	2,006
20	KOCAELI	0,673	0,914	0,905	1,000	1,718	1,576
21	PAMUKKALE	0,267	0,677	0,359	0,679	2,068	0,861
22	TRAKYA	0,427	0,472	0,586	0,542	1,834	1,416
23	SELÇUK	0,682	1,000	1,000	1,000	3,545	0,702
24	ONDOKUZ MAYIS	0,595	0,699	0,902	0,703	1,320	1,036
25	MERSIN	0,461	0,618	0,529	0,653	1,163	2,121
26	MARMARA	1,000	1,000	1,000	1,000	5,598	1,521
27	İNÖNÜ	0,529	0,563	0,698	0,634	2,367	2,373
28	GAZİ	0,595	0,692	1,000	1,000	2,911	0,375
29	EGE.	0,413	0,598	0,482	0,897	1,748	1,060
30	FIRAT	0,379	1,000	0,392	0,993	5,339	0,936
31	YILDIZ TEKNİK	1,000	0,725	1,000	0,823	1,904	2,022
32	YÜZÜNCÜ YIL	0,492	0,608	0,503	0,679	2,560	1,344
33	İSTANBUL	1,000	0,721	1,000	0,743	2,567	4,630
34	AKDENIZ	0,367	0,531	0,372	0,553	1,047	1,757
35	ANKARA	0,633	0,522	1,000	0,702	2,229	1,532
36	SAKARYA	0,994	1,000	1,000	1,000	3,831	2,490
37	ULUDAĞ	0,559	0,608	1,000	0,671	1,392	1,323
	MEAN	0.60	0.74				
	Std.Deviation	0.24	0.18				

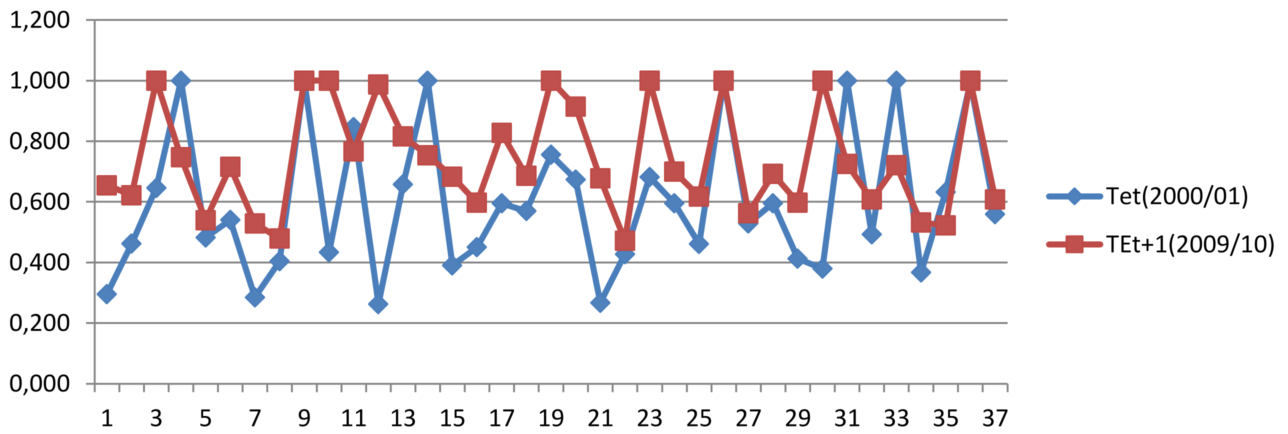


Figure 1. Relative efficiency of 37 public universities in Turkey for periods 2000/01 and 2009/10

The catch-up term compares the closeness of DMU_0 in each period to that period's frontier. A value of one indicates that the DMU_0 has retained its position relative to the frontier in period's *t* and *t+1*. A value above one has the meaning that the DMU_0 has become more efficient in period *t+1* compared to period *t*. Conversely, a value below one implies that the DMU_0 has experienced loss in efficiency. In this context, the first column of Table 2 summarizes this situation. It is obvious that 78% of the universities have experienced growth in efficiency during these two periods.

Boundary shift column in Table 2 shows the measure of the contribution to productivity change of whatever technical change occurs between periods *t* and *t+1* [15]. A value over one indicates a productivity gain by the industry not necessarily by the DMU_0 itself. It states that at the input-output mixes of DMU_0 in periods *t* and *t+1* *efficient production* uses lower input levels in period *t+1* than in period *t* while controlling the output levels [21]. Conversely a value below one indicates productivity loss by the industry. And a value of one implies that there is either a gain or loss in productivity in the industry. On average there is an 18% productivity gain in the higher education sector in Turkey that can be attributable to the boundary shift or technological change in the industry.

Table 2. Efficiency, technological change and productivity growth of Turkish universities.

DMU	University	TEC(effch) (catch-up)	Boundary Shift	Pure Technical Efficiency catchup	Scale Efficiency(t)	Scale Efficiency(t+1)	Scale Efficiency Cathup	TFP
1	ADNAN MENDERES	2,215	0,618	1,463	0,640	0,969	1,514	1,369
2	ATATÜRK	1,345	1,220	0,758	0,540	0,958	1,774	1,641
3	BALIKESİR	1,547	1,407	1,211	0,783	1,000	1,278	2,177
4	CELAL BAYAR	0,747	0,678	0,761	1,000	0,981	0,981	0,506
5	CUMHURİYET	1,117	0,799	0,925	0,807	0,975	1,208	0,892
6	ÇUKUROVA	1,324	1,164	1,098	0,792	0,955	1,206	1,530
7	DİCLE	1,855	1,135	1,980	0,966	0,905	0,937	2,106
8	DOKUZ EYLÜL	1,186	1,356	0,932	0,702	0,893	1,272	1,608
9	DUMLUPINAR	1,000	1,992	1,000	1,000	1,000	1,000	1,992
10	ERCİYES	2,304	0,948	1,795	0,779	1,000	1,283	2,183
11	BOĞAZİÇİ	0,905	1,229	0,938	0,945	0,912	0,966	1,112
12	GALATASARAY	3,758	1,228	1,000	0,263	0,987	3,758	4,613
13	GAZİANTEP	1,242	0,703	1,377	0,915	0,825	0,902	0,873
14	GAZİOSMANPAŞA	0,754	0,564	0,853	1,000	0,883	0,883	0,425
15	HACETTEPE.	1,751	0,756	1,259	0,492	0,683	1,390	1,323
16	HARRAN	1,326	1,023	1,205	0,732	0,806	1,101	1,356
17	İSTANBUL	1,390	2,497	1,000	0,595	0,827	1,390	3,345
18	KARADENİZ TEKNİK	1,203	1,106	0,795	0,622	0,942	1,514	1,331
19	ODTÜ	1,324	1,251	1,319	0,996	1,000	1,004	1,656
20	KOCAELI	1,357	0,896	1,105	0,744	0,914	1,228	1,217
21	PAMUKKALE	2,536	0,973	1,891	0,744	0,997	1,341	2,468
22	TRAKYA	1,103	1,083	0,925	0,729	0,870	1,193	1,196
23	SELÇUK	1,466	1,856	1,000	0,682	1,000	1,466	2,721
24	ONDOKUZ MAYIS	1,176	1,041	0,779	0,660	0,995	1,509	1,224
25	MERSIN	1,341	0,640	1,234	0,871	0,946	1,086	0,858
26	MARMARA	1,000	1,918	1,000	1,000	1,000	1,000	1,918
27	İNÖNÜ	1,063	0,969	0,908	0,758	0,887	1,171	1,030
28	GAZİ	1,163	2,597	1,000	0,595	0,692	1,163	3,004
29	EGE.	1,449	1,067	1,861	0,856	0,666	0,778	1,545
30	FIRAT	2,636	1,501	2,533	0,968	1,007	1,041	3,879
31	YILDIZ TEKNİK	0,725	1,145	0,823	1,000	0,881	0,881	0,826
32	YÜZÜNCÜ YIL	1,234	1,242	1,350	0,979	0,895	0,914	1,533
33	İSTANBUL	0,721	0,879	0,743	1,000	0,970	0,970	0,632
34	AKDENIZ	1,450	0,641	1,487	0,986	0,961	0,975	0,929
35	ANKARA	0,826	1,327	0,702	0,633	0,744	1,176	1,096
36	SAKARYA	1,006	1,237	1,000	0,994	1,000	1,006	1,244
37	ULUDAĞ	1,089	0,983	0,671	0,559	0,907	1,622	1,070
	MEAN	1,396	1,180	1,154				1,633

Pure technical efficiency catch-up is interpreted as the efficiency catch-up. The difference stems from the technology (variable returns to scale) assumed while measuring the efficiency score. And the scale efficiency catch-up represents the impact of any change in scale size of DMU_0 on its productivity.

The last column in Table 2 shows the productivity growth of each university for the periods. As it can be observed from Table 2 78% of the universities have productivity gain. And the average productivity growth for the periods mentioned is %63.

5. CONCLUDING REMARKS

The purpose of the current study is to investigate regress or progress in productivity growth of Turkish universities over the period 2000/01 and 2009/10 by using a DEA-based Malmquist productivity index. The computed results revealed a rise of 63% productivity growth in higher education institutions in Turkey.

References

[1] A. García-Aracil, D. Palomares-Montero, Methodological Problems to Measure University Efficiency in Relation with its Geographic Localization

[2] Bonaccorsi, A., Daraio, C., (2005) Scientometrics, Vol. 63, No. 1

[3] Brennan, S., Haelermans, C., Ruggiero, J., (2014) Nonparametric estimation of education [1] productivity incorporating nondiscretionary inputs with an application to Dutch schools, 234(3),809-818

[4] Casu, B., Girardone, C., Molyneux, P., (2004) Productivity Change in European banking: A comparison of parametric and non-parametric approaches, Journal of Banking & Finance, 28, 2521-2540

[5] Caves, D.W., Christensen, L.R., Diwert, W.E., (1982) The economic theory of index numbers and the measurement of input, output and productivity, Econometrica, 50, 1393-414

[6] Charnes, A., Cooper, W.W., Rhodes, E., (1978), Measuring the efficiency of decision making units, European Journal of Operational Research, 2(6), 391-465

[7] Chen, Y., (2003), A non-radial Malmquist productivity index with an illustrative application to Chinese major industries, Int. J. of Production Economics, 83, 27-35

[8] Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battase, G.E., (2005) An Introduction to Efficiency and Productivity Analysis, Second Edition, Springer

[9] Fare, R., Grosskopf, S., Lindgren, B., Roos, P., (1994) Productivity Developments in Swedish Hospitals: A Malmquist Output Index Approach, Data Envelopment Analysis: Theory, Methodology, and Applications, Springer

[10] Fare, R., Grosskopf, S., Lindgren, B., Roos, P., (1992)Productivity Changes in Swedish Pharamacies 1980-1989: A Non-Parametric Malmquist Approach, The Journal of Productivity Analysis, 3, 85-101

[11] Fare, R., Grosskopf, S., Lovell, C.A.K., (1985) The Measurement of Efficiency of Production, Kluwer Academic Publishers

[12] Fare, R., Grosskopf, S., Norris, M., Zhang, Z., (1994) Productivity Growth, Technical Progress, and Efficiency Change in Industrialized Countries, The American Economic Review, 84 (1)

[13] Flegg, A.T., Allen, D.O., Field, K., Thurlow, T.W., (2004) Measuring the Efficiency of British Universities: A Multi-period Data Envelopment Analysis, Education Economics, 3 (3) 231-249

[14] Forsund, F.R., Edvardsen, D.F., Kittelsen, S.A.C., (2015) Productivity of tax offices in Norway, Journal of Productivity Analysis, DOI 10.1007/s11123-015-0435-1

[15] Grifell-Tatje E., Lovell C.A.K., (1997) The sources of productivity change in Spanish banking, European Journal of Operations Research, 98, 364-380

[16] Kao, C., Liu, S. (2014) Multi-period efficiency measurement in data envelopment analysis: The case of Taiwanese commercial banks, Omega 47,90-98

[17] Lin,Y., Hsu, G.J.Y., Hsiao, C., (2007) Measuring efficiency of domestic banks in Taiwan: application of data envelopment analysis and Malmquist index, Applied Economic Letters, 14, 821-827

[18] Malmquist,S., (1953), Index Numbers and difference surfaces, Trabajos de Estadistica 4,209-232

[19] Sentürk, S.S., (2010) Total Factor Productivity Growth in Turkish Manufacturing Industries: A Malmquist Productivity Index Approach, Master of Science Thesis, Stockholm

[20] Taskin, F., Zaim, O., (1997) Catching-up and innovation in high-and low-income countries, Economic Letters,54,93-100

[21] Thassoulis, E., (2001) Introduction to the Theory and Application of Data Envelopment Analysis: A foundation text with integrated software, Kluwer Academic Publishers

This page blank left intentionally

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2016.04.01.ECON.01

COMPARISON OF COINTEGRATION TESTS FOR NEAR INTEGRATED TIME SERIES DATA WITH STRUCTURAL BREAK

Esin FİRUZAN*

Berhan ÇOBAN †

Faculty of Sciences, Dokuz Eylul University, Izmir, Turkey

Received: 17 December 2015
Accepted: 07 June 2016

Abstract

Sample size of data, presence of structural break, location and magnitude of potential break, and having with near integrated process might affect the performance of cointegration tests. Engle-Granger (EG) and Johansen Cointegration tests may have erroneous results since they do not take into account possible structural break unlike Gregory – Hansen (GH) cointegration test. In this study, it is argued that the suitable choice of cointegration tests is quite complex, since outcomes of these tests are very sensitive to specifying these properties.

The performance of cointegration tests is compared to each other underlying properties. This study presents how standard residual based tests- Engle-Granger and Gregory-Hansen- for cointegration can be implemented if series is near integrated, that is close to a unit root process. For assessing the finite sample performance of these tests, a Monte-Carlo experiment showed that both cointegration tests have relatively better size and power properties depend on break point, break magnitude, sample size of time series and the hypothesized value of AR(1) parameter. To illustrate the findings of the paper, a financial data is analyzed. The practitioners should be careful about the hypothesized value of AR(1) parameter which represents dependency degree of the data. If the autoregressive parameters is very close to one and the break magnitude is high, any test is acceptable for moderate to large sample size. However, one might need very large sample size to have a good power and actual size of the test. Additionally, GH test becomes liberal test unlike EG test as the magnitude of structural break increases..

Keywords: Cointegration, Structural Break, Engle- Granger Test, Gregory-Hansen Test

YAPISAL KIRILMALI İÇ BAĞIMLILIĞI YÜKSEK ZAMAN SERİLERİNDE EŞBÜTÜNLEŞME TESTLERİNİN KARŞILAŞTIRILMASI

Özet

Örneklem büyüklüğü, yapısal kırılmanın varlığı, potansiyel kırılmanın yeri ve büyüklüğü ve birim köke yakın prosese sahip olmak Eşbütünleşme testlerinin performanslarını etkileyebilir. Engle-Granger (EG) ve Johansen eşbütünleşme testleri, Gregory – Hansen (GH) eşbütünleşme testinden farklı olarak, olası kırılmaları dikkate almadığından hatalı sonuçlar verebilmektedir. Sözü geçen testlerin çıktıları bu özelliklerin yapısına çok duyarlı olduğundan, bu çalışmada uygun eşbütünleşme testinin seçilmesinin oldukça karmaşık olduğu tartışılmıştır.

Eşbütünleşme testlerinin performansları belirtilen özellikler altında karşılaştırıldı. Bu çalışma, standart hata terimi tabanlı testlerin - Engle-Granger ve Gregory-Hansen- serilerin yüksek iç bağımlılığa (birim köke yakın süreçlere) sahip olduğunda nasıl uygulanabileceğini göstermektedir. Testlerin sonlu örneklem performansları değerlendirildiğinde, Monte Carlo deney sonuçları, her iki testin de kırılma noktası, kırılmanın büyüklüğü, serinin genişliği ve AR(1) parametresi değerleri için anlamlılık düzeyi ve güç değerleri açısından iyi sonuçlar verdiğini göstermiştir. Çalışmanın bulguları finansal veri ile de analiz edilmiştir. Araştırmacılar AR(1) modelin iç bağımlılığını gösteren

* esin.firuzan@deu.edu.tr (Corresponding author)
† berhan.coban@deu.edu.tr

parametrenin değerini test ederken dikkatli olmalıdırlar. Otoregresif modelin parametresinin bire çok yakın çıktığı ve yapısal kırılmanın büyüklüğünün yüksek olduğu durumda her iki test de büyük örneklem genişliği altında uygulanabilir. Ancak testlerin daha iyi güç değerlerine ve nominal anlamlılık düzeylerine sahip olması için çok büyük örneklemlere ihtiyaç vardır. Ek olarak yapısal kırılmanın büyüklüğü arttıkça Gregory – Hansen testi Engle Granger testine göre daha liberal davranışlar sergilemektedir.

Anahtar Kelimeler : Eşbütünleşme, Yapısal Kırılma, Engle- Granger Testi, Gregory-Hansen Testi

1. INTRODUCTION

Most of the economic variables such as inflation, interest rates, exchange rates, real GDP and so forth appear to be highly persistent, and are typically described as unit root or nonstationary processes. Structural break/change in the series sometimes may cause the unit root or nonstationarity problems. Time series with structural changes and unit roots share similar features that make it difficult to discriminate between the two fundamentally nonstationary processes. A well-known drawback of the conventional unit root test is their potential confusion of structural changes in the series as evidence of unit root. Misleading inferences in unit root process may also cause misleading inferences in cointegration process in ignoring an existing structural break. They may potentially fail to reject the cointegration hypothesis if the series have a structural break. In other words, there may be a possibility that the series are in fact cointegrated around the structural break(s), however they are mistakenly classified by noncointegrated. Starting with pioneer study of Perron(1989), Hendry and Neale(1991) and Lee *et al*(1997) examine the performance of testing the null of stationarity when structural break is ignored. Ignoring an existing structural break can give rise to apparent unit roots in stationary time series. Kwiatkowski et al. (1992) propose stationarity tests that take into account of structural breaks. Following this development, many authors, including El-Shagi and Giesen (2013), Zivot and Andrews (1992) and Perron (1997) proposed determining the break point 'endogenously' from the data.

Many cointegration tests have been developed in the literature in order to avoid inference problems in misleading decisions in cointegration under structural break. The most widely applied cointegration tests are residual-based ones for the null hypothesis of no cointegration is tested in the sense of Engle & Granger (1987) two-step method. According to Engle & Granger, two time series are cointegrated if a linear combination of the integrated series (I(1)) has a stationary distribution. Gregory- Hansen (1996) and Campos et al. (1996) propose cointegration tests that allow for a structural break of unknown timing. Harris and Inder(1994), Hao (1996), Bartley, Lee and Strazicich (2001), and Carrion-I-Silvestre and Sanso (2006) develop test that examine for the null of cointegration with one structural break in the level and slope. These developed cointegration tests rely upon the strict unit-root assumption which is not easy to argue on economics theory. Although many economic variables are frequently modeled as unit root processes, there is a little a priori information that these variables have an exact unit root, rather than a root close to unity. Since unit-root tests have low power to distinguish between a unit-root and a close to unity under the structural break(s), many econometricians and finance researchers started to discuss the possibility of mild unit root assumption instead of strictly unit root assumption. Phillips (1988) considers near-integrated process that has roots smaller than unity (strongly autoregressive) in his analysis. De Boef & Granato (1999) argue that cointegration tests may lead analyst to conclude mistakenly that the data are cointegrated (or near-cointegrated) when the data are near-integrated. Hjalmarsson & Österholm (2007) investigate the properties of Johansen's maximum eigenvalue and trace tests for cointegration under the empirically relevant situation of near-integrated variables.

This study contributes to the empirical literature by further re-examining the size and power of the EG and GH tests when the data have deviations from the strictly unit-root assumption because of structural break on the determination of near-cointegrated or noncointegrated. It is assumed that cointegration relationship can be modeled by level shift, level shift in trend and regime shift by near integrated time series. The sensitivity of EG and GH tests to sample size and unity conditions of data is examined and how well these tests can discriminate under the conditions in respect of the study undertaken.

The paper is organized as follows. Section 2 makes a brief description of the EG and GH cointegration tests. Section 3 explains how to construct the experimental design of the study for revealing the impact of break location, break magnitude and dependency degree parameter. Also, results of the Monte Carlo experiment in terms of finite sample size and power are discussed in this section. It is tried to make a clear whether they are conservative or liberal tests. Section 4 gives the results of earthquake and tsunami impacts on Tokyo and Chinese Stock Markets. It is clear that the earthquake has caused a structural change in both markets. The application is performed gradually depending on the magnitude and location of structural break. Finally, some concluding remarks are offered in the conclusion.

2. COINTEGRATION TESTS

Cointegration methodology is the innovation used in time series econometrics in the last decade, by stating the possible existence of long-run equilibrium relationships among nonstationary series (I(1)). Two or more time series are said to be cointegrated if each of the series are individually nonstationary while some linear combination of the series is stationary (I(0)). Cointegration means that many shocks can cause permanent changes in the I(1) series but there is some long-run equilibrium relation. The equilibrium can be distorted in the short run but not in the long run.

Given the linear cointegrating regression model,

$$y_t = \alpha' c_t + \beta' x_t + u_t \qquad (1)$$

where $z_t = (y_t, x_t)'$ is a *m*-vector I(1) time series which may or may not have cointegrating relationship. c_t is a vector of deterministic terms (such as a constant or time trends). $\{y_t\}$ and $\{x_t\}$ are cointegrated if $\{u_t\}$ is stationary. Brief description of the cointegration tests that take into account of structural break is provided in this section.

2.1. Engle-Granger Test (EG)

Cointegration analysis begins with pretests for unit root in the individual series of interest. In brief, the series of interest is first- differenced and regressed on its own lagged levels. If the coefficient on lagged levels is statistically significant different from zero, the data are best characterized as I(1) processes. Since first-differenced series is stationary, the Engle-Granger (1987) methodology can be conducted at two stages. Levels of dependent process are linearly regressed on levels of independent process. The residuals from this regression are also tested for the stationarity using unit root tests. Rejection of the null of residual series obtained from the cointegrating regression has unit root, gives strong evidence of cointegration. For the statistical theory and overview, see Granger (1981, 1983) Engle-Granger (1987), Johansen (1995).

2.2. Gregory-Hansen Test (GH)

Another commonly used test is Gregory-Hansen test (GH), which consider an alternative hypothesis in which the cointegrating vector may be subject to a structural break at an unknown/known time. Gregory – Hansen (1996) analyzed models that arrange, under the alternative hypothesis of cointegration, the possibility of changes in parameters. GH test investigates the determination of the structural breaks in the cointegration analysis in three models such as a break in the level, a break in level in the presence of a linear trend and a break in both level and linear trend.

These models are:

2.2.1. Level shift model (C)

This model, which involves a break in the level, is expressed as below**:**

$$y_t = \mu_1 + \mu_2 DU_t + \beta' x_t + u_t \qquad t = 1,2,\dots,T \qquad (2)$$

μ_1 represents the constant term before the break, μ_2 represents the change in the constant term during the structural break, β' indicates the cointegrating slope coefficient, DU_t denotes a dummy variable which allows break to occur at time τ such that $DU_t = 1$ if $t > [T\tau]$ and zero otherwise. Here *n* is the number of observations, τ is a coefficient which denotes the break period between (0.15*T*, 0.85*T*).

2.2.2. Level shift with trend (C/T)

Level shift with trend model which involves the break in intercept with trend is expressed as below:

$$y_t = \mu_1 + \mu_2 DU_t + \alpha t + \beta' x_t + u_t \quad t = 1,2,\dots,T \quad (3)$$

Model in equation 3 has same parameters as in equation 2, but additionally includes trend variable as t for measuring trend effect.

2.2.3. Regime shift (C/S)

Regime shift model which involves a break in both the intercept and trend (slope of trend line) of the series is expressed as below.

$$y_t = \mu_1 + \mu_2 DU_t + \beta_1' x_t + \beta_2' x_t DU_t + u_t$$
$$t = 1,2,\dots,T \qquad (4)$$

μ_1 , μ_2 and DU_t have the same representation as in level shift model. β_1' represents the cointegrating slope coefficient before the break, β_2' denotes the change in the slope coefficient after the break. Model C/S is different from Model C/T since the former involves trend variable. GH test tries to reveal the cointegrated structure depending on the location of the break.

3. FINITE-SAMPLE EVIDENCE-MONTE CARLO EXPERIMENT

Data generation process (DGP) is conducted using MATLAB (R2011a). The series are generated for three different models which consist of single structural break in intercept, break in intercept and trend and break in intercept and slope of the time series model.

Cointegrated y_t and x_t series are generated from AR(1) and I(1) process with near-integrated values of ϕ=0.9,0.95,0.99 parameter for 50, 100, 200 sample size with 10000 replications. In general, the closer ϕ is to one; the less likely the series are cointegrated. Behavior of u_t series which are produced by equation 1 under consideration have examined depend on stationarity condition. If it is stationary, y_t and x_t series are cointegrated.

In this study, it is assumed that the break date is known or exogenous. The power of the tests is compared depending on part of the series where break has been occurred. The parts of the series are treated as break in the first quarter (0.25T), in the second quarter (0.50T) and in the third quarter (0.75T) of the series. Since the magnitude of the structural break in the series may also affect the power of the tests, the performance of tests is also investigated with 1, 5 and 10 unit break magnitude in corresponding parts of the series.

3.1. Power&Emprical Size Comparison of the Cointegration Tests for All Models

To evaluate the performances of cointegration tests and to compare them to each other, we carried out a simulation separately, for each model. Totally eighty figures, graphics of power and empirical size, are examined in the experiment, but substantial figures are presented in the following sections.

3.1.1. Level shift model

If ϕ is close to unity (ϕ=0.99), then increasing in break magnitude in small sample size 50 leads to decrease power for both test. Impact of structural break which is occurred especially in the last quarter has negative affect both tests' power, but power of EG test has fallen down dramatically in comparison with GH. Increasing the size of the sample from 50 to 100 is not enough to have high power level for each test. To attain such a high level of power, increasing sample size 200 rather than 100 is a good solution for the relatively low levels of ϕ coefficient (for 0.9 and 0.95). However, if ϕ is 0.99, both tests' powers have decreased deeply. Remarkably point is that the shape of the GH test's power is saddle-shaped when ϕ is 0.95 as shown in Figure 1.

In comparing of tests based on empirical size, the empirical size of EG test is increasing if the structural break is in the first quarter and last quarter part of the series. Also, it reaches to 9% level as sample size increases. The empirical size approaches the nominal size as structural break magnitude increases and if it is in the middle of the series, in contrast to GH test, whose empirical size level is at around %2 as sample size and structural break magnitude increase. GH test is liberal in this model under the consideration.

3.1.2. Level shift with trend model

Since trended time series is used in this model, even if sample series of length is 50, increasing ϕ coefficient (almost have a unit root), does not cause any decline rapidly in both tests' power. As magnitude of structural break goes up, so does power of GH test becomes larger than the ones of EG test. However, there is a point at which increasing sample size (from 50 to 200) couldn't suffice to keep the power 80% level especially when ϕ=0.99. The results are shown in Figure 2. As sample size increases, the ϕ coefficient becomes more important for the power of these tests in this model. The power of EG test has negatively affected by magnitude and location (if it is in the middle of the series) of structural break and closeness of ϕ coefficient to unit root whereas power of GH test has negatively affected by just location of the structural break.

According to this model, EG test has the lowest value of empirical size when sample size is 50. As sample size increases, location and magnitude of structural break have no longer significant impacts on the empirical size which has grown to almost 6% level. Growing empirical size means that the discrepancy of empirical size and nominal size is getting bigger. In contrast to EG test, the empirical size of GH test approaches to nominal size under sample size increment regardless of location and magnitude of structural break.

3.1.3. Regime shift with decreasing trend

The series which is generated by simulation consists of decreasing trend component in this model. Because of trend effect, both tests' power is decreasing if the break is especially in the first quarter part of the series as sample size increases. But this decreasing in power becomes dramatic for EG test. It would need a much larger sample size in order to increase the power. The results could be seem in Figure 3.

When behaviors of empirical size of these tests are examined for this model, empirical size of EG test drifts away from the nominal size as the number of sample size increases whereas one of GH test approaches to nominal size. Empirical size of GH test does not affected by location and magnitude of structural break under the regime shift with decreasing trend model.

3.1.4. Regime shift with increasing trend

Both tests have high level power performance (100%) in regime shift with increasing trend model as shown in Figure 4. Power of EG test has been affected by dependency degree of the series (ϕ coefficient) as sample size increases. Besides the magnitude of structural break, the location of the break becomes crucial in this model. The EG test's power is at the deep level if the break has occurred in the last quarter of the series. Increasing sample size makes the GH test more sensitive to location and magnitude of break. The GH test has low power when structural break occurred in the last quarter of the series with 10 unit increment. Increasing the dependency degree of the series from 0.9 to 0.99 does not change this consequences.

Figure 1. Level Shift Model

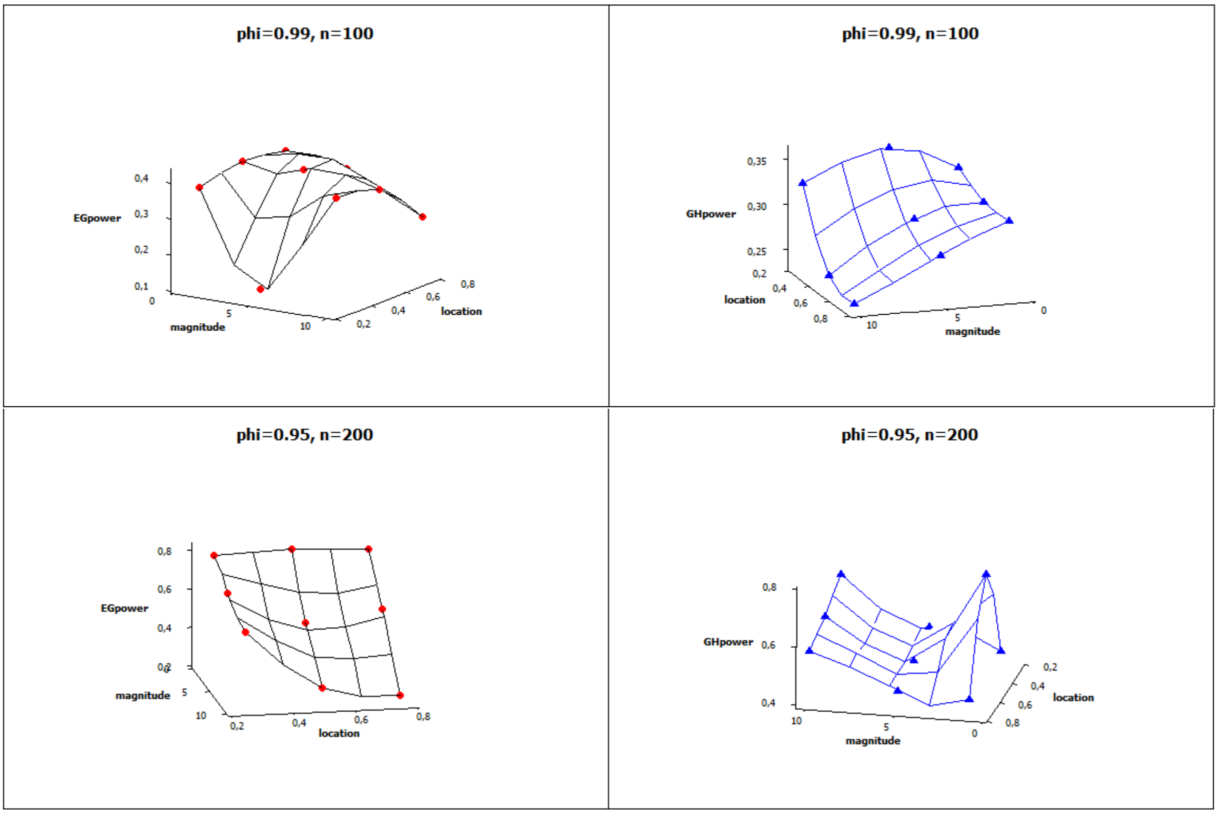


Figure 2. Level Shift with Trend Model

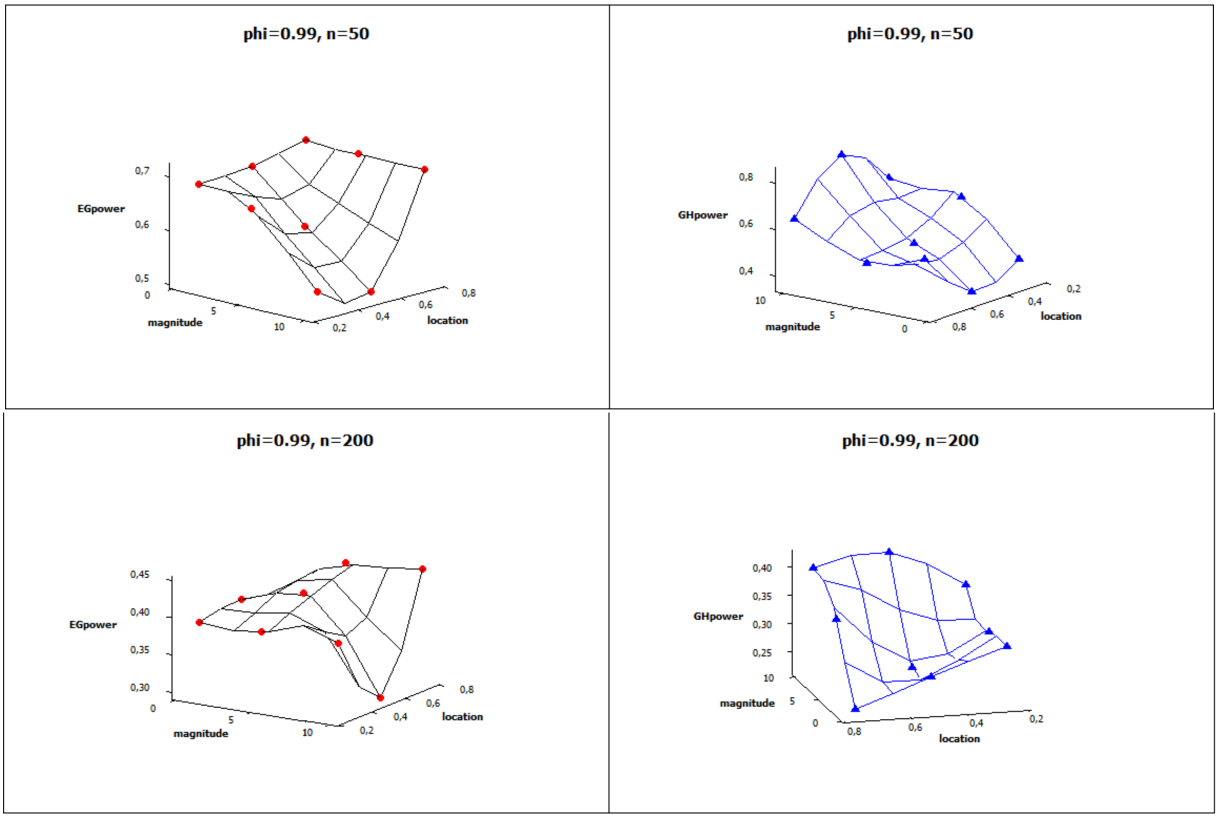


Figure 3. Regime Shift Model (with Decreasing Trend after structural break)

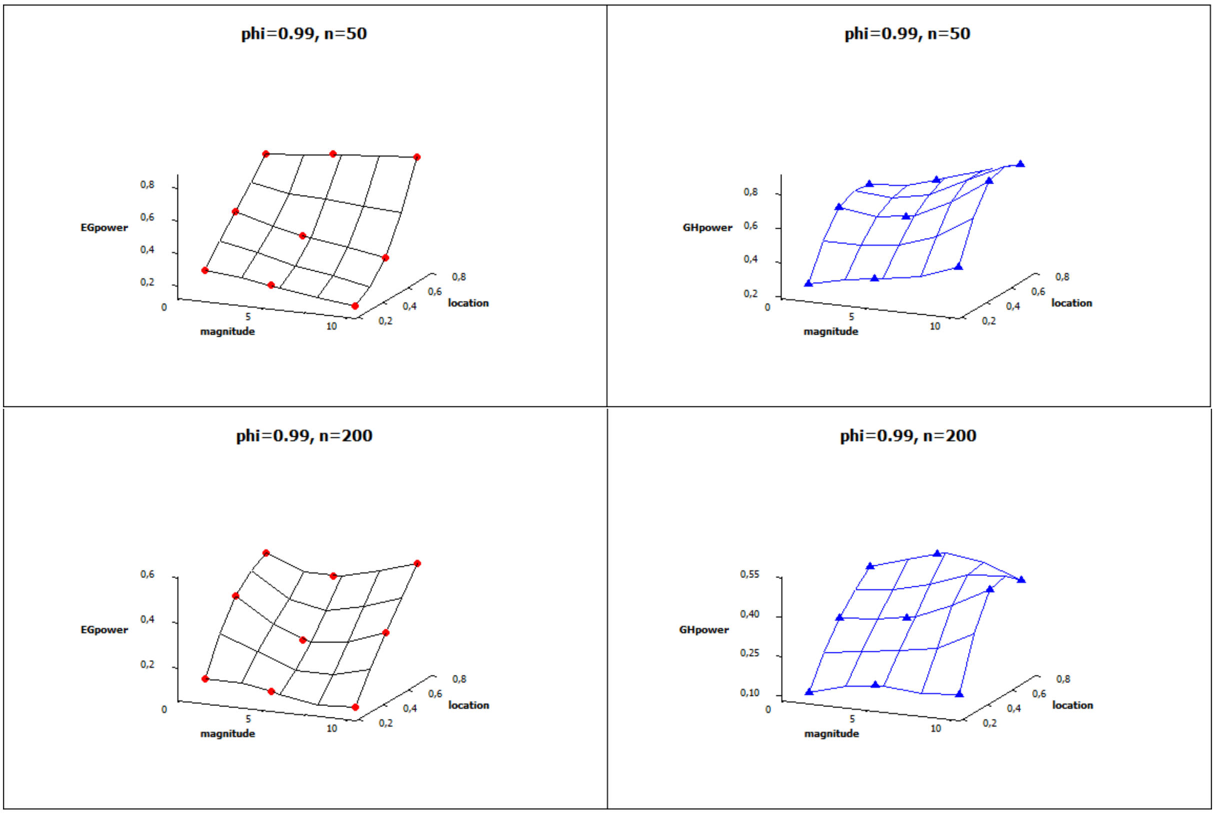


Figure 4. Regime Shift Model (with Increasing Trend after structural break)

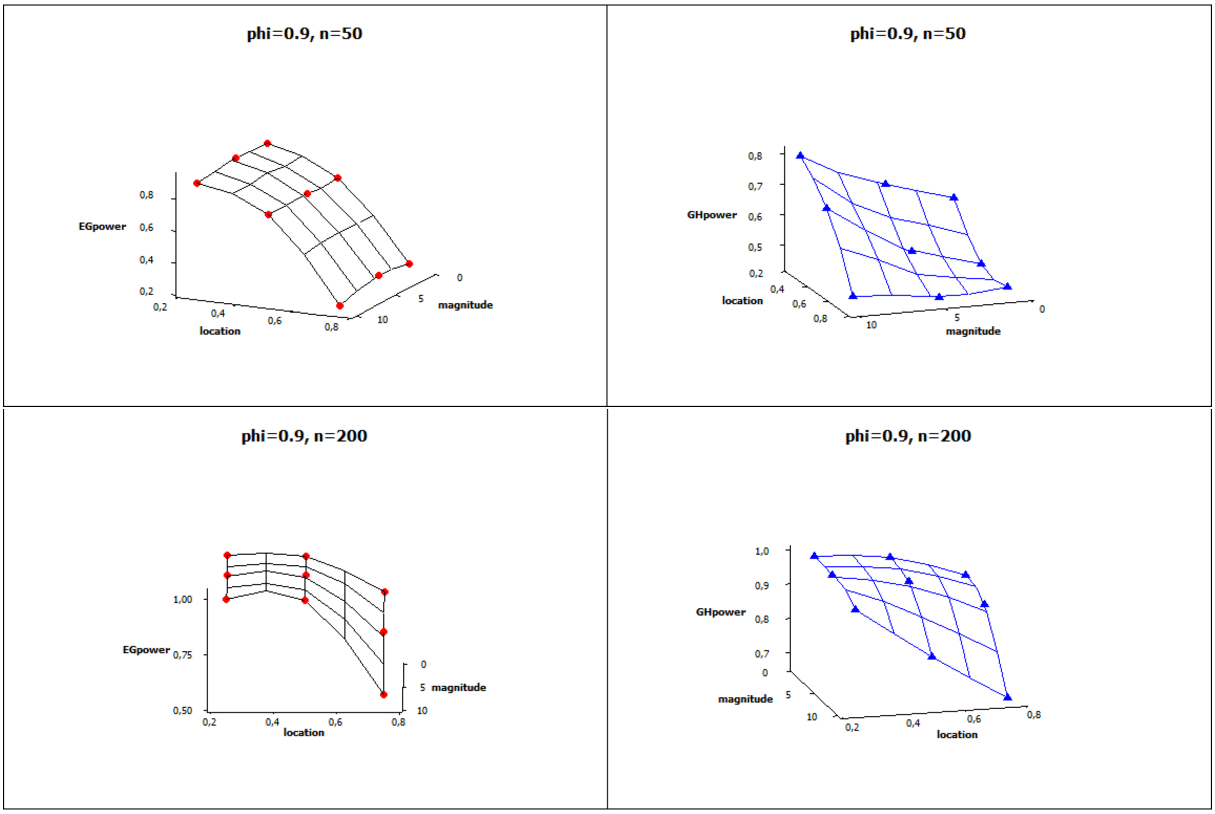
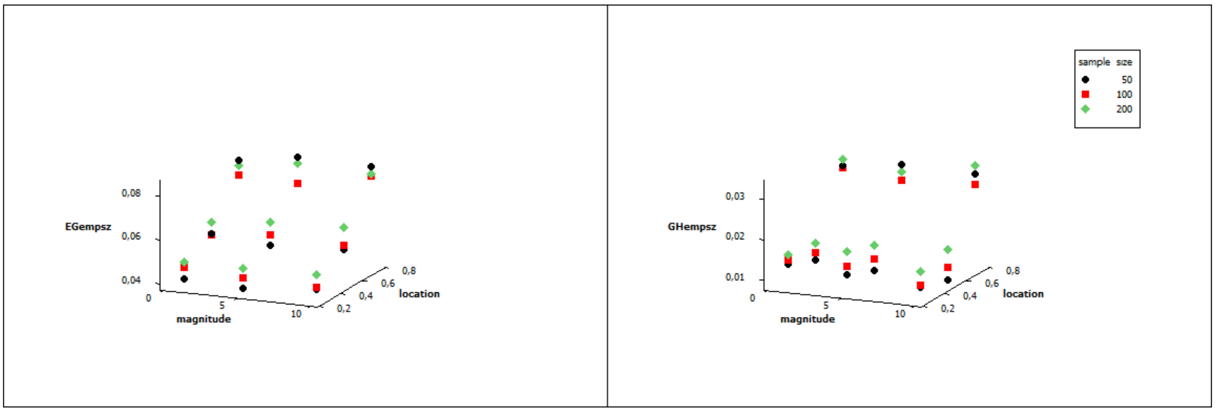


Figure 5. Empirical size comparison of the tests for regime shift model (with increasing trend)

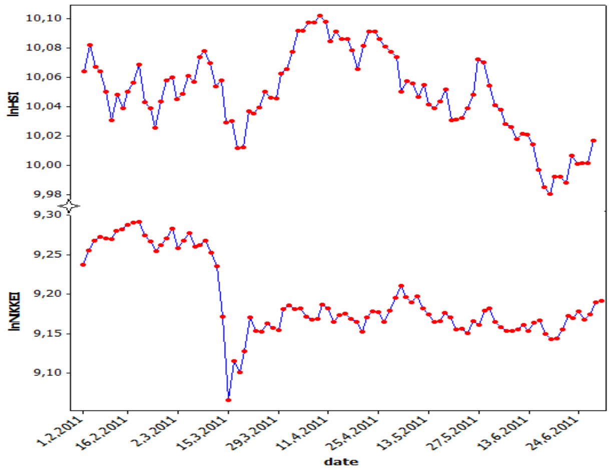


Empirical size comparisons of the tests for regime shift model are shown in Figure 5. According to Figure 5, getting larger sample size does not make any difference in tendency of empirical size. According to location of structural break becomes important for EG test. If the break has occurred in the first quarter part of the series, empirical size of EG test becomes very close to nominal size. Reasonable agreement between empirical and nominal size is relevant for GH test, but the location of the break. If the break has occurred in the last quarter part of the series, empirical size of GH test approaches to nominal size.

4. EMPIRICAL ILLUSTRATION

To illustrate what is stressed in this study that it must be taken account of not only location but also magnitude of structural break before performing cointegration tests (EG and GH), Tokyo Stock Market and Chinese Stock Market, which is thought cointegrated with Tokyo stock market, daily data have been used in this paper. The data covers a period of six month period starting from Feb.1st, 2011 and ending in Jan. 30th, 2011. The data of these market time series are plotted in the graphs at level in Figure 6.

Figure 6. Time series plot of earthquake impact on Tokyo and Chinese stock markets.



The negative impact of the March 2011 earthquake and tsunami on Tokyo Stock market has been particularly felt on the 15th of March which is the second working day after the devastating earthquake in Japan. The NIKKEI dropped over 10% to finish at 8605.15, a loss of 1015 points. Japan's massive earthquake and tsunami have sent shockwaves through China's economy. Japan was one of China's most important trade partners, accounting for about 8 percent of China's total exports, while China sourced 13 percent of its imports from Japan. Trade between the two countries was definitely be affected in the short term. Although the quake also reduced Japans' imports from China, overall the spike was relatively modest and short-lived. The later and weaker impact has been showed up in HSI stock market.

To evaluate the influence of the location of structural break, the data is examined as separating into three sets: The first set contains the structural break which is in the first quarter part of the series (T_1=94), the second one covers the structural break in the middle of the series (T_2=50) and the third one is included it in the last part of the series (T_3=37). The data is analyzed into three models (level shift, level shift with trend and regime shift). According to Zivot Andrews unit root test, the structural break has occurred in the 15th of March. The EG test results showed that the series are not cointegrated regardless of location of the structural break. [the first quarter part (p-value=0.6556), middle (p-value=0.8471) and the last part of the series (p-value=0.8638)]. According to GH test results, the series are cointegrated if the structural break has occurred in the first quarter part (with ϕ=0.941), in the middle part (with ϕ=0.932) and in the last quarter part of the series (with ϕ=0.852) p-value<0.001, p-value<0.001, p-value<0.025), respectively.

5. CONCLUSION

In this study, effects of systematic pattern of time series data on cointegration tests is investigated. This study revealed that in the presence of structural break, the specification of the systematic pattern is also so important in performance of EG and GH tests. It has been shown that in the presence of a structural break, these tests' conclusion about cointegrating relationship can be biased toward not rejecting it when the near integrated time series data is used. In a finite sample it is quite difficult to distinguish between an integrated variable and a "near-integrated" variable. Therefore, if a priori knowledge of a structural break exists is evident in eyesight, one should use the GH cointegration test.

The magnitude of the structural break is also important for choosing the cointegration tests. If there is a slight one unit drop in the series (i.e. drop from 100 to 99), researchers can choose both tests for analysis, because the drop does not cause significant effect on the series. However, when the drop is around 10 unit, the effect of structural break must not ignore, then the GH cointegration test must be preferred.

Identifying the model type (level shift, level shift with trend and regime shift) is the also important part of cointegration analysis especially with near-integrated data, because existing of trend component in the series may block to distinguish the structural break. In this circumstance, there need to take larger sample size.

The increasing of dependency parameter (ϕ) have negatively impact on level shift model. In level shift model with trend, EG test results has more sensitivity to the location and high dependency degree, whereas GH test has affected by just the location of the structural break. When the series fits to regime shift model with increasing trend, the location of structural break become considerable for both test, but the power of EG test is decreasing deeply when the break has occurred in the last part of the series.

GH test becomes liberal test, so its power does not quickly decline unlike EG test as the magnitude of structural break increases

References

Akdi, Y. (2003), “Zaman serileri analizi (birim kökler ve kointegrasyon)”, Ankara. Bıçaklar Kitabevi

Bartley, W. A., Lee, J., Strazicich, M. C., (2001), “Testing the null of cointegration in the presence of a structural break”, Economics Letters 73 (3), 315–323.

Campos J., Ericsson N.R. and Hendry.D.F.(1996),”Cointegration test in the presence of structural breaks”, Journal of Econometrics, 70, 187–220.

Carrion-i-Silvestre, J. L., Sans`o, A., (2006). “Testing the null of cointegration with structural breaks.”, Oxford Bulletin of Economics and Statistics 68 (5), 623–646.

Enders, W. (1995). “Applied econometric time series (2nd ed.)”,Iowa State University. John Wiley & Sons.Inc.

Engle, R.F., Granger C. W. J. (1987). "Cointegration and Error Correction: Representation, Estimation and Testing", Econometrica 55,251-276

El-Shagi, M., Giesen S. (2013). "Testing for structural breaks at unknown time: A Steeplechase", Comput. Econ., 41,101-123

Gregory, A. W., Hansen B. E. (1996),” Residual-based tests for cointegration in models with regime shifts”, Journal of Econometrics, 70, 99-126

Granger, C.W.J. (1981), ”Some Properties of Time Series Data and Their Use in Econometric Model Specification”, Journal of Econometrics, 16, 121-130.

Granger C.W.J (1983): "Co-Integrated Variables and Error-Correcting Models," unpublished UCSD Discussion Paper 83-13.

Hamilton, J. D. (1994),” Time series analysis”, New Jersey: Princeton University Press.

Harris, D., Inder, B. (1994). “A test of the null of cointegration”, In: Hargreaves, C., ed. Nonstationary Time Series Analysis and Cointegration. Oxford: Oxford University Press, pp. 133–152.

Hao, K. (1996), “Testing for structural change in cointegrated regression models: some comparisons and generalizations”, Economet. Rev. 15:401–429.

Hendry, D.F., A.J. Neale (1991), “A Monte Carlo Study of the Effects of Structural Breaks on Tests for Unit Roots”, in P.Hackl and A. Westlund (eds.), Economic Structural Change, Springer Verlag, New York, 95-119

Hjalmarsson E., Österholm P.(2007). “A residual-based cointegration test for near root variables”, International Finance Discussion Papers with number 907.

Johansen S. (1995). "Likelihood-Based Inference in Cointegrated Vector Autoregressive Models", Oxford. Oxford University Press

Kadılar, C. (2000), "Uygulamalı Çok Değişkenli Zaman Serileri Analizi", Ankara: Bizim Büro Basımevi

Kwiatowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992) "Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root", Journal of Econometrics, 54, 159–78.

Lee, J., Huang, C. J. and Shin, Y. (1997) "On stationarity tests in the presence of structural breaks", Economics Letters, 55, 165–72.

Maddala, G. S. (1992). "Introduction to econometrics (2nd ed.)", New York: Macmillan Publishing Company

Maddala G. S. & Kim I.M. (2003), "Unit roots, cointegration and structural change (5th ed.)", Cambridge: Cambridge University Press.

Perron, P.(1989). "The great crash, the oil-price shock, and the unit-root hypothesis. Econometrica", Vol. 57, No. 6. (Nov., 1989), pp. 1361-1401

Perron, P. (1997), "Further Evidence on Breaking Trend Functions in Macroeconomic Variables", Journal of Econometrics, 80 (2), pp.355-385.

Phillips, P. C. B. (1988b), "Regression theory for near-integrated time series", Econometrica, 56, 1021-1043

Suzanna D.B., Granato J. (1999). 'Testing for cointegrating relationshipswith near-integrated data.' Political Analysis, 8(1): 99–117.

Yurdakul, F. (2000). Yapısal kırılmaların varlığı durumunda geliştirilen birim-kök testleri. Gazi Üniversitesi İ.İ.B.F. Dergisi, 2, 21-34

Zivot, E., & Andrews D.W.K. (1992). Further evidence on the great crash, the oil- price shock, and the unit-root hypothesis. Journal of Business & Economic Statistics 10 (3), 251-270.

This page blank left intentionally

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2016.04.01.STAT.02

UNSHARED AND SHARED FRAILTY MODELS

Nihal ATA TUTKUN *

Assoc. Prof. Dr., Hacettepe University, Ankara

Diren YEĞEN †

Republic of Turkey, Ministry of Culture and Tourism

Received: 28 December 2016
Accepted: 27 April 2016

Abstract

The Cox regression model which is commonly used in survival analysis is established under the proportional hazards assumption. However cases in which the data shows heterogeneity come across in studies. In this case, heterogeneity should be explained in order to make the interpretations more effective which were obtained depending on the model. Frailty models are one of the survival analysis methods which were developed for explaining heterogeneity.

In this study, frailty models are examined theoretically and were applied to the lung cancer data. The unshared frailty model has been used to explain the difference between general risk and momentary risk of individuals in the data set. As for comparing the momentary risk between individuals with various levels of explanatory variables with other individuals, shared frailty models have been used.

Keywords: Survival analysis, Cox regression, nonproportional hazards, parametric regression models, frailty models

PAYLAŞILMAMIŞ VE PAYLAŞILMIŞ ZAYIFLIK MODELLERİ

Özet

Yaşam çözümlemesinde sıklıkla kullanılan Cox regresyon modeli orantılı tehlikeler varsayımı altında kurulmaktadır. Ancak çalışmalarda verinin heterojen özellik gösterdiği durumlar ile karşılaşılmaktadır. Bu durumda modele bağlı olarak elde edilen yorumların daha etkin olabilmesi için heterojenliğin açıklanması gerekmektedir. Zayıflık modelleri heterojenliğin açıklanması için geliştirilmiş bir yaşam çözümlemesi yöntemidir.

Bu çalışmada, zayıflık modelleri teorik açıdan incelenmiş ve akciğer kanseri verisi kullanılarak bir uygulama yapılmıştır. Veri kümesindeki bireylerin taşıdığı genel risk ile herhangi bir bireyin anlık riski arasındaki farklılığı açıklamada paylaşılmamış zayıflık modeli kullanılmıştır. Açıklayıcı değişkenlerin çeşitli düzeylerine sahip bireylerin veri kümesindeki diğer bireylere göre anlık riskinin karşılaştırılmasında ise paylaşılmış zayıflık modelleri kullanılmıştır.

Anahtar Kelimeler : Yaşam çözümlemesi, orantısız tehlikeler, parametrik regresyon modelleri, zayıflık modelleri

* nihalata@hacetepe.edu.tr (Corresponding author)
† dy11163@hotmail.com

1. GİRİŞ

Yaşam çözümlemesi mühendislik, tıp, biyoloji ve demografi gibi bilim dallarında kullanılan temel bir araştırma yöntemidir. Yaşam çözümlemesi, tıbbi ve demografik çalışmalarda incelenen ölümlülük kavramının bir karşılığı olarak ortaya çıkmıştır. İlgilenilen olayın ortaya çıkma süresine yani başarısızlık süresi verilerine dayanan bu araştırma yöntemi Cox (1974)'un geliştirdiği Orantılı Tehlikeler Modeli yaklaşımıyla beraber geniş bir uygulama alanına yayılmıştır.

Yaşam çözümlemesinde, bağımlı değişken olarak ele alınan yaşam süresinin açıklayıcı değişkenler tarafından etkilenebileceği göz önünde bulundurulduğunda, regresyon modellerinin yaşam çözümlemesinde önemli bir yere sahip olduğu görülmektedir. Bu modellerden biri de zayıflık (frailty) modelidir. Zayıflık modeli özellikle tıp, biyoloji ve genetik çalışmalarının da içinde bulunduğu çeşitli alanlarda kullanılmaktadır. Orantılı tehlikeler modelinde aynı değişken değerine sahip olan birimlerin aynı yaşam süresine sahip olacağı varsayılmaktadır. Ancak bu durum, aynı tedavi, yaş ve cinsiyet grubundaki gözlenen tüm bireylerin aynı gözlenen yaşam sürelerine sahip olduğu anlamına gelmemektedir. Bazı çalışmalarda, ölçülen açıklayıcı değişkenler dışında yaşam süresini önemli derecede etkileyen ancak gözlenemeyen başka faktörler de olabilir. Bu durum, birimlerin heterojenliği olarak belirtilmektedir. Zayıflık modelinin temeli, birimler arasındaki heterojenliği açıklamak için ölçülemeyen rasgele etkiyi modele dahil etmektir.

Bu çalışmanın amacı, zayıflık modellerinin yapısını ve türlerini incelemektir. İkinci bölümde, zayıflık modellerinde kullanılan yöntemler verilmiştir. Üçüncü bölümde ise akciğer kanseri verilerine klasik yaşam çözümlemesi yöntemlerinin yanı sıra zayıflık modelleri uygulanmış ve elde edilen sonuçlar incelenmiştir.

2. Zayıflık Modelleri

Yaşam çözümlemesinde zamana bağlı verilerin bağımsız ve aynı dağılımdan geldiği yani kitlenin homojen olduğu varsayılmaktadır. Fakat gözlemler incelenince birimlerin aynı dağılımdan gelmediği, bağımsız olmadığı aksine kendine özgü özellikleri olduğu söylenebilir. Bu farklılık ile kitle heterojen bir nitelik kazanmaktadır. Bu heterojenliği değerlendirmek zordur, ancak önemlidir. Heterojenlik başlıca şu iki nedenden kaynaklanabilir:

Gözlenebilen risk faktörlerinden kaynaklanan değişkenlik,

Bilinmeyen açıklayıcı değişkenlerden kaynaklanan heterojenlik.

Heterojenlik orantısız veya azalan tehlikeler gibi bazı beklenmedik sonuçları açıklamaktadır. Eğer bazı birimler yüksek başarısızlık riski taşıyorsa, diğerleri daha az riskli bir grup oluşturmaya eğilimlidirler.

Gözlenemeyen zayıflığı göz önüne almadan tahmin edilen bireysel tehlike oranı, zaman geçtikçe tehlike fonksiyonunda yapısal bir bozulmaya yol açmaktadır. Kitlenin farklı riskler taşıyan birimlerin karışımı olduğu varsayılırsa karma modeller kullanılabilir. Gözlenemeyen riskler zayıflık olarak tanımlanmaktadır. Bu zayıflık değeri bilinemediğinden tehlike fonksiyonuna çarpımsal olarak dahil edilmektedir. Çünkü birimler ile kitle arasındaki ilişkinin yapısı, zayıflığın birimler arasındaki dağılımına bağlıdır. Zayıflık terimi farklı dağılım türlerine uyabilir. Hougaard (1984, 1995), Clayton (1978), Yashin vd. (1995), Oakes (1967), Congdon (1995) çalışmalarında zayıflık terimi için en çok kullanılan dağılımlar, Gamma ve ters-Gauss dağılımlarıdır. Zayıflık dağılımının varyansı ise incelenen kitledeki heterojenlik derecesini belirlemektedir.

Zayıflık modeli ilk kez Vaupel (1979) tarafından mortalite çalışmalarında uygulanmıştır. Lancaster (1979) işsizlik sürelerinin modellenmesinde zayıflık modelini kullanmıştır. Zayıflık modeli ile ilgili çalışmalar birçok araştırmacının ilgisini çekmiştir (McGilchrist ve Aisbett, 1991; Guo ve Rodriquez, 1992, Yashin and Iachine, 1995; Babiker ve Cuzick, 1994; Hougard, 2000; Wienke ve diğ., 2000, 2004; O'Quigley ve Stare, 2002, 2004). Bu konuyla ilgili çalışmalar özellikle biyoloji ve genetik çalışmalarının da içinde bulunduğu çeşitli bilim dallarında kullanılmıştır.

Zayıflık modelleri paylaşılmış ve paylaşılmamış zayıflık modelleri olmak üzere başlıca iki bölümde incelenmektedir.

2.1. Paylaşılmamış Zayıflık Modeli

Paylaşılmamış zayıflık modelleri başarısızlık süresi ilişkisiz olan birimleri incelemektedir. Bu yaklaşımda kitlenin, belli gözlenebilir değişkenler ile açıklandığı ve homojen olduğu varsayımı dikkate alınmaktadır. Örneğin, tıbbi bir çalışmada bireylerin farklı ilaçlara ya da uygulanan tedaviye farklı tepkiler verdiği gözlenebilir. Heterojenlik açıklanması zor fakat bir o kadar da gerekli bir durumdur. Yapılan çalışmalarda, zayıflık modelleri bu heterojenliği açıklamak adına kullanılmaktadır. Bu yaklaşımların temeli, bireylerin farklı zayıflıklara sahip olduğu ve en zayıf olanın, daha az zayıf olandan daha erken başarısızlığa uğrayacağıdır.

En çok uygulanan zayıflık modeli orantılı tehlikeler modelini öngörmektedir. Bu da rasgele etki üzerinde koşullanmıştır. Birimlere ait tehlike fonksiyonu; gözlenemeyen, zamandan bağımsız rasgele değişken olan Z'ye bağlıdır. Bu rasgele değişken temel tehlike fonksiyonuna çarpımsal olarak etki eder ve

$$h(t,Z) = Zh_0(t) \quad (1)$$

biçimindedir. Burada Z kitle içinde değişkenlik gösteren rasgele değişken olarak ele alınır. Zayıflık modelleri E(Z)=1 ve V(Z)=σ^2 olacak biçimde standartlaştırılmıştır. Burada Z'nin varyansı temel riskin içindeki heterojenliğin ölçüsü olarak yorumlanabilir. Eğer Z=1 olursa zayıflık dağılımını içeren model orantılı tehlikeler modeline dönüşmektedir.

Eşitlik 1'de verilen çarpımsal zayıflık modeline göre, zayıflık zamandan bağımsız ve temel tehlike fonksiyonuna çarpımsal olarak etki etmektedir. Bu durum da yaşam çözümlemesindeki gözlenemeyen heterojenliğe neden olmaktadır. Eşitlik 1'e bilinen açıklayıcı değişkenler eklenirse:

$$h(t,Z,X) = Zh_0(t)\exp(\beta^T X) \quad (2)$$

elde edilir. Burada X=(X_1,...,X_k) açıklayıcı değişkenler vektörü ve β=(β_1,...,β_k) regresyon parametreleri vektörüdür. $S(t/Z)$, zayıflık (Z) kısıtı altında birimlerin yaşam fonksiyonunu göstermektedir ve

$$S(t/Z) = \exp(-\int_0^t h(s,Z)ds) = \exp(-Z\int_0^t h_0(s)ds) = \exp(-ZH_0(t))$$

biçiminde elde edilir. Burada $H_0(t) = \int_0^t h_0(s)ds$ temel bikimli tehlike fonksiyonudur.

Kitlenin yaşam fonksiyonu, yaşam fonksiyonlarının ortalamasıdır. Yaşam fonksiyonu ve olasılık yoğunluk fonksiyonu zayıflık fonksiyonunun ortalaması ve varyansı ile nitelendirilir. Olabilirlik fonksiyonu ise $(T_i, \Delta_i, X_i, Z_i)$ ve $(i = 1,2,...,n)$ için

$$\prod_{i=1}^{n} (Z_i h_0(t_i)\exp(\beta^T X_i))^{\delta_i} \exp(-Z_i h_0(t_i)\exp(\beta^T X_i))$$

biçiminde verilmektedir.

Zayıflığın her bir birim için sürecin tamamı boyunca belirlendiği varsayılmaktadır. Fakat kitle bileşenleri süreç boyunca değişikliğe uğrar. Zayıf birimler erken başarısızlığa eğilimlidir ve ilk başarısız olacak olanlardır. Buna bağlı olarak risk altındaki kitlenin zayıflık dağılımı süreç ilerledikçe değişmektedir.

Bu nedenle kitleye ait ortalama zayıflık $\int_0^\infty z\, f(z \backslash T > t)dz$, zamanla azalacaktır. Aşağıdaki teorem bu yapıyı açıklamaktadır.

Teorem: Eşitlik 1 ile verilen model gözönüne alınırsa; $h(t) = \dfrac{f(t)}{S(t)}$ eşitliğinden yararlanarak,

h(t)=E(h(t,Z)\T>t) elde edilebilir. Daha açık biçimiyle;

$$h(t) = \int_0^\infty h(t,Z) f\left\langle z \middle| T > t \right\rangle dz = h_0(t)\int_0^\infty zf\left\langle z \middle| T > t\right\rangle dz$$

olur. Burada, f(z\T>t) t anında hayatta kalanlara ait zayıflığın yoğunluk fonksiyonudur (Wienke, 2011).

2.2. Zayıflık Terimi için Kullanılan Dağılımlar

Uygulamalarda en çok rastlanan zayıflık terimi dağılımları; Gamma ve Ters-Gauss dağılımlarıdır.

2.2.1. Gamma Zayıflık Modeli

Gamma dağılımı uygulamalarda sıklıkla karşılaşılan dağılımlardan biridir. Gamma dağılımı yaşam verilerine iyi uyum sağlayan bir dağılımdır. Gamma dağılımının Laplace dönüşümü aracılığıyla, birikimli yoğunluk fonksiyonu, tehlike fonksiyonu rahatça açıklanabilmektedir. Gamma dağılımı parametreleri sayesinde $\Gamma(k,\lambda)$, k=1 olduğunda üstel dağılıma, k çok büyük seçildiğinde normal dağılıma benzer olması bu dağılımı kullanışlı biçime getirmektedir. Vaupel, Manton ve Stallard (1979) geliştirdikleri yaklaşımla zayıflık terimi $Z \sim \Gamma(k,\lambda)$ olduğunda modelin yapısını ortaya koymuştur. k konum, λ ölçek parametresi olmak üzere Z'nin marjinal yoğunluk fonksiyonu,

$$f(z) = \frac{\lambda^k z^{k-1} e^{-\lambda z}}{\Gamma(k)}; \qquad z > 0, k > 0, \lambda > 0$$

olarak tanımlanmaktadır. Laplace dönüşümü,

$$L(u) = \frac{1}{\Gamma(k)}\lambda^k \int e^{-uz} z^{k-1} e^{-\lambda z} dz$$

$$= \frac{\lambda^k}{(\lambda+u)^k}\frac{1}{\Gamma(k)}(\lambda+u)^k \int z^{k-1} e^{-(\lambda+u)z} dz$$

biçimindedir. k ve (λ+u) parametreleriyle gamma dağılımının Laplace dönüşümünün birinci ve ikinci türevi,

$$L'(u) = -\frac{k}{\lambda}(1+\frac{u}{\lambda})^{-k-1},$$
$$L''(u) = \frac{k(k+1)}{\lambda^2}(1+\frac{u}{\lambda})^{-k-2}$$

olarak verilmektedir. u=0 noktasında bu eşitlikler yardımıyla beklenen değer ve varyans,

$$E(Z) = \frac{k}{\lambda}, \quad V(Z) = \frac{k(k+1)}{\lambda^2} - \frac{k^2}{\lambda^2} = \frac{k}{\lambda^2}$$

biçimde elde edilmektedir. k=λ kısıtı altında bu dağılım E(Z)=1 ve $\sigma^2 = \frac{1}{\lambda}$ olmaktadır.

$Z \sim \Gamma(\frac{1}{\sigma^2}, \frac{1}{\sigma^2})$ olan gamma dağılımının olasılık yoğunluk fonksiyonu;

$$f(z) = \frac{1}{\Gamma(\frac{1}{\sigma^2})}(\frac{1}{\sigma^2})^{\frac{1}{\sigma^2}} z^{\frac{1}{\sigma^2}-1} \exp(-\frac{z}{\sigma^2})$$

olmakta ve koşulsuz yaşam fonksiyonu, koşulsuz olasılık yoğunluk fonksiyonu ve tehlike fonksiyonu ise sırasıyla aşağıda verilmektedir.

$$S(t) = L(H_0(t)) = \frac{1}{(1+\sigma^2 H_0(t))^{\frac{1}{\sigma^2}}}$$

$$f(t) = \frac{h_0(t)}{(1+\sigma^2 H_0(t))^{\frac{1}{\sigma^2}+1}},$$

$$h(t) = \frac{h_0(t)}{1+\sigma^2 H_0(t)}.$$

Yaşayan bireyler/birimler üzerinden zayıflık dağılımının olasılık yoğunluk fonksiyonu,

$$f(z|X, T>t) = \frac{S(t|X,z) f(z)}{S(t|X)}$$
$$= \exp(-zH_0(t)e^{\beta^T X}) z^{\frac{1}{\sigma^2}-1} \exp(-z(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}))$$
$$= \frac{(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X})^{\frac{1}{\sigma^2}}}{\Gamma(\frac{1}{\sigma^2})} z^{\frac{1}{\sigma^2}-1} \exp(-z(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}))$$

olarak tanımlanmaktadır. Benzer biçimde t anında başarısızlığa uğrayanlar için olasılık yoğunluk fonksiyonu aşağıdaki biçimde verilmektedir:

$$f(z|X, T=t)$$
$$= \frac{f(t|X,z) f(z)}{f(t|X)}$$
$$= \frac{zh_0(t)\exp(-zH_0(t)e^{\beta^T X}) z^{\frac{1}{\sigma^2}-1} \exp(-\frac{z}{\sigma^2})}{\Gamma(\frac{1}{\sigma^2})\sigma^{\frac{2}{\sigma^2}} h_0(t)(1+\sigma^2 H_0(t)e^{\beta^T X})^{-\frac{1}{\sigma^2}-1}}$$
$$= \frac{(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X})^{\frac{1}{\sigma^2}+1}}{\Gamma(\frac{1}{\sigma^2}+1)} z^{\frac{1}{\sigma^2}+1-1} \exp(-z(\frac{1}{\sigma^2} + H_0(t)e^{\beta^T X}))$$

Yukarıdaki eşitlikten yararlanarak t anındaki ölümlerin ortalaması;

$$E(Z|X, T=t) = \frac{1+\sigma^2}{1+\sigma^2 H_0(t)e^{\beta^T X}}$$

ve t anından sonraki ölümlerin ortalaması;

$$E(Z|X, T>t) = \frac{1}{1+\sigma^2 H_0(t)e^{\beta^T X}}$$

olarak verilmektedir. t anında başarısız olan birimler için zayıflık daha yüksek bir ortalama vermektedir. Bu durum yüksek risk taşıyan birimlerin daha erken başarısızlığa uğrayacağının bir işaretidir. t anında başarısız olan birimlerin varyansı,

$$V(Z|X,T=t)=\frac{\sigma^2(1+\sigma^2)}{(1+\sigma^2 H_0(t)e^{\beta^T X})^2}$$

iken yaşayanlar için varyans;

$$V(Z|X,T>t)=\frac{\sigma^2}{(1+\sigma^2 H_0(t)e^{\beta^T X})^2}$$

olarak tanımlanmaktadır. Bu iki eşitlikten de görüldüğü gibi zayıflığın varyansı zamanla azalmaktadır (Gutierrez, 2002; Wienke, 2011).

2.2.2. Ters-Gauss Zayıflık Modeli

Ters-Gauss zayıflık modeli, gamma zayıflık modelinin bir alternatifi olarak önerilmiştir. Bu zayıflık modeli Hougaard (1974) tarafından geliştirilmiş Klein (1992), Keiding, Andersen ve Klein (1997), Price ve Manatunga (2001), Economou ve Caroni (2005), Kheiri (2007), Duchateau ve Janssen (2007) tarafından da incelenmiştir.

$\mu>0$ ve $\lambda>0$ olmak üzere ters-Gauss dağılımın olasılık yoğunluk fonksiyonu;

$$f(z)=\frac{\sqrt{\lambda}}{\sqrt{2\pi z^3}}\exp(-\frac{\lambda}{2\mu^2 z}(z-\mu)^2)$$

olarak verilmektedir. Ters-Gauss dağılımın Laplace dönüşümü ise aşağıdaki biçimdedir:

$$L(u)=E(e^{-uZ})=\exp\left(-\frac{\lambda\sqrt{1+\frac{2\mu^2u}{\lambda}}}{\mu}+\frac{\lambda}{\mu}\right).$$

$$=\exp\left(\frac{\lambda}{\mu}(1-\sqrt{1+\frac{2\mu^2u}{\lambda}}\right)$$

Laplace dönüşümlerinin birinci ve ikinci türevleri yardımı ile beklenen değer ve varyans hesaplanabilmektedir. Birinci ve ikinci türevler, sırasıyla,

$$L'(u)=-\frac{\mu}{\sqrt{1+\frac{2\mu^2u}{\lambda}}}\exp\left(\frac{\lambda}{\mu}(1-\sqrt{1+\frac{2\mu^2u}{\lambda}})\right),$$

$$L''(u)=\frac{\mu^3}{\lambda(1+\frac{2\mu^2u}{\lambda})^{\frac{3}{2}}}\exp\left(\frac{\lambda}{\mu}(1-\sqrt{1+\frac{2\mu^2u}{\lambda}})\right)$$
$$+\frac{\mu^2}{1+\frac{2\mu^2u}{\lambda}}\exp\left(\frac{\lambda}{\mu}(1-\sqrt{1+\frac{2\mu^2u}{\lambda}}\right)$$

olarak verilmektedir. Beklenen değer ve varyans ise u=0 noktasındaki türevlerin hesaplanması ile aşağıdaki biçimde ifade edilmektedir:

$$E(Z)=-L'(0)=\mu,$$
$$V(Z)=L''(0)-(L'(0))^2=\frac{\mu^3}{\lambda}.$$

Eğer $E(Z)=\mu=1$ ve $V(Z)=\sigma^2=\frac{1}{\lambda}$ alınırsa Laplace dönüşümü, $L(u)=e^{\frac{1}{\sigma^2}(1-\sqrt{1+2\sigma^2u})}$ olmaktadır. Bu dönüşüm sayesinde yaşam fonksiyonu ve tehlike fonksiyonu elde edilebilmektedir. Yaşam fonksiyonu ve tehlike fonksiyonu sırasıyla,

$$S(t)=e^{\frac{1}{\sigma^2}(1-\sqrt{1+2\sigma^2H_0(t)})}$$

$$h(t)=\frac{h_0(t)}{(1+2\sigma^2H_0(t))^{\frac{1}{2}}}$$

biçimindedir. t anından itibaren yaşayan birimlere ait zayıflığın olasılık yoğunluk fonksiyonu,

$$f(z|X,T>t)=\frac{S(t|X,z)f(z)}{s(t|X)}$$
$$=\frac{1}{\sqrt{2\pi^2z^3}}\exp\left(-\frac{(z-(1+2\sigma^2H_0(t)e^{\beta^TX})^{-\frac{1}{2}})^2}{\frac{2\sigma^2z}{1+2\sigma^2H_0(t)e^{\beta^TX}}}\right)$$

olarak tanımlanmaktadır. Bu birimler için beklenen değer

$E(Z|X,T>t)=\frac{1}{\sqrt{1+\sigma^2 H_0(t)e^{\beta^T X}}}$ ve varyans ise

$V(Z|X,T>t)=\frac{\sigma^2}{(1+\sigma^2 H_0(t)e^{\beta^T X})}$ biçimindedir

(Wienke, 2011).

2.3. Paylaşılmış Zayıflık Modeli

Paylaşılmış zayıflık modeli ile ilgili ilk çalışmalar Clayton (1978) ve Clayton and Cuzick (1985) tarafından yapılmıştır. Hougaard (1986) Weibull bireysel tehlike fonksiyonu ile paylaşılmış zayıflık modelini, Whitmore ve Lee (1991) üstel bireysel tehlike fonksiyonu ile ters-Gauss paylaşılmış zayıflık modelini ve Sahu (1997) ise Gibbs örneklemesini kullanarak Bayesci paylaşılmış zayıflık modelini incelemişlerdir. Xue ve Brookmeyer (1996) paylaşılmış zayıflık modelinin kısıtları üzerine çalışma yapmıştır. Ibrahim ve diğ. (2001) parametrik modeller gibi yarı parametrik paylaşılmış zayıflık modellerine de Bayesci yaklaşım uygulamıştır. Klein ve Moeschberger (2003) yarı parametrik paylaşılmış zayıflık modeline EM algoritmasını uygulamışlardır.

Modelin paylaşılmış zayıflık modeli olarak adlandırılmasının nedeni gruptaki birimlerin aynı zayıflığı paylaşmasıdır (Clayton, 1978). Bu birimlerin başarısızlık süreleri koşullu bağımsızdır. Bu koşul, zayıflık (Z) üzerinden şekillenmektedir. Her birimin sahip olduğu temel tehlike fonksiyonu $Zh_0(t)$ biçimindedir. Birimlerin başarısızlık süreleri arasındaki bağımlılık durumu, Z değerinin birimler için ortak olmasından kaynaklanmaktadır. Bu yaklaşımda zayıflık terimi olayların oluş zamanları arasındaki ilişkiyi modellemede kullanılmaktadır. Bu yaklaşıma göre, paylaşılmış zayıflık modeli, benzer kümelerde bulunan birimler için ortak bir zayıflık terimine sahiptir. Bu zayıflık (Z) rasgele dağılmaktadır. Yani, bu model koşullu bağımsız bir modeldir ve zayıflık grup içindeki tüm birimler için ortaktır. Dolayısıyla olayların oluş zamanları arasında bir bağlılık yaratır.

Bu yaklaşımdaki tehlike modeli aynı paylaşılmamış zayıflık modelindeki gibidir. En önemli fark paylaşılmış zayıflık modelinde, zayıflık birimlerin grup içerisinde paylaştığı ilişkili riski ifade etmektedir. Bu yüzden zayıflık birimler yerine birimlerin oluşturduğu grubu ifade etmektedir.

Paylaşılmış zayıflık yaklaşımına göre, gruptaki tüm başarısızlık sürelerini verilen zayıflıktan koşullu bağımsızdır. Zayıflık teriminin değeri zaman boyunca sürekli ve kümedeki birimler için ortaktır. Bu durum kümedeki başarısızlık süreleri üzerindeki bağımlılığın nedenidir. Bu bağımlılık paylaşılmış zayıflık modellerinde hep pozitif değerler almaktadır.

Yaşam çözümlemesinde paylaşılmış zayıflık modeli, n adet küme ve i ($1 \le i \le n$) olmak üzere kümesinin n_i adet gözleme sahip olduğu varsayımı altında; gözlemlerin Z_i ($1 \le i \le n$) zayıflıklarına sahip olduğu ifade edilmektedir. ($1 \le i \le n, 1 \le j \le n_i$) olmak üzere X_{ij} ise i'inci kümedeki j'inci gözlemin başarısızlık süresi olan T_{ij} hakkında bilgi içeren açıklayıcı değişkenler vektörünü ifade etmektedir. Her bir i ($1 \le i \le n$) kümesindeki yaşam süreleri, Z_i üzerinden koşullu olarak tehlike fonksiyonlarından bağımsız olarak varsayılmaktadır.

Paylaşılmış zayıflık modeli,

$$h(t/X_{ij},Z_i)=Z_i h_0(t)\exp(\beta^T X_{ij}) \tag{3}$$

biçiminde verilmektedir. Burada $h_0(t)$ temel tehlike fonksiyonunu ve β tahmin edilebilen parametreler vektörünü göstermektedir. Zayıflıklar (Z_i), birbirinden bağımsız ve aynı dağılımlı ve θ zayıflık dağılımının parametresi olmak üzere aynı f(z,θ) olasılık yoğunluk fonksiyonuna sahiptir.

Çok değişkenli yaşam fonksiyonu zayıflık (Z_i) üzerinden koşullu olmak üzere, i. gruptaki birimler için,

$$\begin{aligned} &S(t_{i1},...,t_{in_i}/X_i,Z_i)\\ &=S(t_{i1}/X_{i1},Z_1)...S(t_{in_i}/X_{in_i},Z_i)\\ &=\exp(-Z_i\sum_{j=1}^{n_i}H_0(t_{ij})\exp(\beta^T X_{ij})) \end{aligned} \tag{4}$$

biçimdedir. $H_0(t)=\int_0^t h_0(s)ds$ ve

$X_i=(X_{i1},...,X_{in_i})$ açıklayıcı değişkenler matrisi olmak üzere Eşitlik 4'ün Z_i üzerinden beklenen değeri marjinal yaşam fonksiyonunu vermektedir:

$$\begin{aligned} &S(t_{i1},...,t_{in_i}/X_i,Z_i)=\\ &=E\exp\left(-Z_i\sum_{j=1}^{n_i}H_0(t_{ij})\exp(\beta^T X_{ij})\right)\\ &=L\left(\sum_{j=1}^{n_i}H_0(t_{ij})\exp(\beta^T X_{ij})\right). \end{aligned}$$

Burada L zayıflık değişkeninin Laplace dönüşümünü ifade etmektedir. Böylece, çok değişkenli yaşam fonksiyonu temel tehlike fonksiyonunun dağılım fonksiyonunda hesaplanan zayıflık dağılımının Laplace dönüşümü ile açıklanabilir.

Birleşik yaşam fonksiyonu ise yaşam fonksiyonlarından Eşitlik 5 ile elde edilebilir. Buradaki gruplar arasında bağımsızlık olduğu varsayımı bahsi geçen çıkarımı elde etmede kritik noktadır.

$$S(t_{11},...,t_{nn_n} / X_1,...,X_n) = \prod_{i=1}^{n} L\left(\sum_{j=1}^{n_i} H_0(t_{ij})\exp(\beta^T X_{ij})\right). \quad (5)$$

Tek değişkenli koşulsuz yaşam fonksiyonları Laplace dönüşümünün bir sonucu olarak elde edilebilmektedir:

$$S(t_{ij} / X_{ij}) = ES(t_{ij} / X_{ij}, Z_i)$$
$$= E\exp(-Z_i H_0(t_{ij})\exp(\beta^T X_{ij}))$$
$$= L(H_0(t_{ij})\exp(\beta^T X_{ij})).$$

L^{-1} Laplace dönüşümünün (L) tersi olmak üzere,

$$H_0(t_{ij})\exp(\beta^T X_{ij}) = L^{-1}(S(t_{ij} / X_{ij}))$$

ve i. grup için koşulsuz yaşam fonksiyonu,

$$S(t_{i1},...,t_{in_i} / X_i) = L(L^{-1}(S(t_{i1/X_{i1}})) + ... + L^{-1}(S(t_{in_i} / X_{in_i})))$$

biçimindedir. n grup sayısını göstermek üzere her bir gruptaki birim sayısı n_i (j=1,…,ni, i=1,…,n) olmak üzere i. gruptaki j. birim için başlangıç zamanı, bitiş zamanı ve başarısız ya da durdurulmuş olması $(t_{0ij}, t_{ij}, d_{ij})$ ile gösterilsin. Bu durumda paylaşılmış zayıflık modelinin olabilirlik fonksiyonu,

$$L_{ij}(\alpha_i) = \frac{S_{ij}(t_{ij} / \alpha_i)}{S_{ij}(t_{0ij} / \alpha_i)}\left[h_{ij}(t_{ij} / \alpha_i)\right]^{d_{ij}}$$
$$= \left[\frac{S_{ij}(t_{ij})}{S_{ij}(t_{0ij})}\right]^{\alpha_i}\left[\alpha_i h_{ij}(t_{ij})\right]^{d_{ij}}$$

biçimindedir. $D_i = \sum_{j=1}^{n_i} d_{ij}$ olursa, i. grup için olabilirlik fonksiyonu

$$L_i(\alpha_i) = \alpha_i^{D_i}\prod_{j=1}^{n_i}\left[\frac{S_{ij}(t_{ij})}{S_{ij}(t_{0ij})}\right]^{\alpha_i}\left[h_{ij}(t_{ij})\right]^{d_{ij}}$$

ifadesinde α_i ye göre integral alınarak,

$L_i = \int_0^\infty L_i(\alpha_i)g(\alpha_i)d\alpha_i$ biçiminde hesaplanır.

Zayıflık terimi Gamma dağılımına sahip ise aşağıdaki gibi hesaplanır:

$$L_i = \left[\prod_{j=1}^{n_i}\left[h_{ij}(t_{ij})\right]^{d_{ij}}\right]\frac{\Gamma(1/\theta + D_i)}{\Gamma(1/\theta)}\theta^{D_i}\left[1 - \theta\sum_{j=1}^{n_i}\ln\frac{S_{ij}(t_{ij})}{S_{ij}(t_{0ij})}\right].$$

Paylaşılmış gamma zayıflık modeli (gözlenen açıklayıcı değişkenler modele dahil edilirse),

$$\hat{Z}_i = \frac{1/\sigma^2 + \sum_{j=1}^{n_i}\delta_{ij}}{1/\sigma^2 + \sum_{j=1}^{n_i} e^{\beta X_{ij}} H(t_{ij})}, \quad i=1,...,n$$

biçiminde yazılır (Gutierrez, 2002; Wienke, 2011).

3. UYGULAMA

3.1. Veri

Bu çalışmada, Ata (2005)'de orantısız tehlikeler için yaşam modellerinin incelendiği çalışmanın verileri ele alınmıştır. İbn-i Sina Hastanesi Göğüs Cerrahisi Bölümü'nde tedavi gören 236 akciğer kanseri hastasına ait veriler kullanılmıştır. Hastalar, ameliyat olduktan sonra hastalıklarının ilk nüks etmesine kadar geçen süre (min=1 ay, max=93 ay) boyunca izlenmiştir. Akciğer kanseri hastalarının yaşam sürelerini etkileyen faktörler yaşam çözümlemesi yöntemleri kullanılarak belirlenmeye çalışılmıştır.

Çalışmada, hastaların ameliyat olduğu tarihten hastalığın ilk nüksetmesine kadar geçen süre (ay olarak) yaşam süresi olarak alınmıştır. Hastalığın nüksetmesi başarısızlık olarak ifade edilmiştir. Hastalığı nüksetmeyen hastalar durdurulmuş olarak tanımlanmıştır. Hastaların izlenme süresi sona erdiğinde 236 hastadan 94'ünde (%39.8) başarısızlık ve 142'sinde (%60.2) durdurma gözlenmiştir. Uygulamada yaş (YŞ), sigara tüketimi (paket yıl olarak, ST), genişletilmiş rezeksiyon (extended resection, ER), tümörün boyutu (mm olarak, BY), invazyon (İV) ve patalojik evre (PE) değişkenleri gruplandırılarak çözümlemeye alınmıştır. Bu değişkenler ve değişkenlerin düzeyleri Tablo 1.'de verilmiştir.

Tablo 1. Kullanılan Değişkenler ve Düzeyleri

Değişken	Değişken Düzeyleri	n	%	Başarısız	D
Yaş	1. <=39	13	5.5	5	8
	2. 40-49	38	16.1	12	26
	3. 50-59	76	32.2	33	43
	4. 60-69	81	34.3	32	49
	5. >=70	28	11.9	12	16
Sigara Tüketimi	1. <=5	16	6.8	4	12
	2. 6-30	62	26.3	23	39
	3. 31-60	123	52.1	48	75
	4. >=61	35	14.8	19	16
Genişletilmiş Rezeksiyon	0. Yok	190	80.5	71	119
	1. Var	46	19.5	23	23
Boyut	1. <=30	73	30.9	25	48
	2. 31-40	46	19.5	12	34
	3. 41-50	41	17.4	18	23
	4. >=50	76	32.2	39	37
İnvazyon	0. Yok	136	57.6	50	86
	1. Var	100	42.4	44	56
Patolojik Evre	1. Evre I	102	43.2	28	74
	2. Evre II	61	25.8	24	37
	3. Evre III	60	25.4	30	30
	4. Evre IV	13	5.5	12	1

D:Durdurulmuş

Bu çalışmada ise hızlandırılmış başarıssızlık zamanı modelleri ve zayıflık modelleri aynı veri kümesi için incelenmiş ve Ata (2005)'in çalışmasında bulunan sonuçlar ile birlikte değerlendirilmiştir.

3.2. Hızlandırılmış Başarısızlık Süresi Modeli Sonuçları

Hızlandırılmış başarısızlık süresi (HBS) modeli gibi parametrik yöntemler yaşam süresi bilinen bir dağılıma uygunluk gösteriyorsa, parametrik olmayan ya da yarı parametrik yöntemlere göre daha iyi sonuçlar vermektedir. Akciğer kanseri veri kümesi için HBS modelleri kapsamında üstel, Weibull, log-lojistik, log-normal, Gompertz, Gamma HBS modelleri ile çözümleme yapılmıştır. Modellerin anlamlılığını test etmek için olabilirlik oranı (LR) test istatistiği kullanılmış ve tüm modellerin istatistiksel olarak anlamlı olduğu görülmüştür($p < 0.05$). İncelenen modeller için AIC değerleri sırasıyla Üstel (AIC=478.250), Weibull (AIC=475.778), Log-lojistik (AIC=467.979), Log-normal (AIC=471.586), Gompertz (AIC=480.214), Gamma (AIC=468.593) olarak elde edilmiştir. Log-lojistik HBS modeli akciğer kanseri verisi için kullanıldığında elde edilen sonuçlar Tablo 2'de verilmiştir.

Tablo 2. Log-lojistik HBS Modeli Çözümlemesinin Sonuçları

Değişken ve düzeyleri		$\hat{\beta}$	S.H.	Alt sınır – Üst sınır	p değeri
Kesişimterimi		5.3139	0.5773	4.1825 – 6.4454	**< 0.0001**
Yaş	1. <=39				
	2. 40-49	0.1174	0.4483	-0.7613 – 0.9962	**0.7934**
	3. 50-59	-0.2057	0.4139	-1.0168 – 0.6054	**0.6191**
	4. 60-69	-0.0162	0.4198	-0.8390 – 0.8066	**0.9692**
	5. >=70	-0.3570	0.4870	-1.3115 – 0.5975	**0.4636**
Sigara Tüketimi	1. <=5				
	2. 6-30	-0.5232	0.4564	-1.4177 – 0.3714	**0.2517**
	3. 31-60	-0.415	0.4330	-1.2501 – 0.4471	**0.3538**
	4. >=61	-1.0662	0.4642	-1.9760 - -0.1563	**0.0216**
Genişletilmiş Rezeksiyon	1. Yok				
	2. Var	-0.2812	0.2552	-0.7815 – 0.2190	**0.2705**
Boyut	1. <=30				
	2. 31-40	0.3124	0.3041	-0.2837 – 0.9084	**0.3043**
	3. 41-50	-0.4470	0.3060	-1.0469 – 0.1528	**0.1441**
	4. >=50	-0.7138	0.2627	-1.2286 – 0.1989	**0.0066**
İnvazyon	1. Yok				
	2. Var	-0.1851	0.2549	-0.6846 – 0.3144	**0.4676**
Patolojik Evre	1. Evre I				
	2. Evre II	-0.1737	0.2893	-0.7408 – 0.3933	**0.5482**
	3. Evre III	-0.6372	0.2634	-11534 – 0.1211	**0.0155**
	4. Evre IV	-1.5326	0.3570	-2.2322 – 0.8330	**< 0.0001**
Ölçek (σ)		**0.6669**	**0.0586**	**0.5614 – 0.7922**	

Yaşam çözümlemesinde kullanılan regresyon tipi modellerde değişken düzeylerinden biri referans kategorisi olarak alınmakta ve değişken düzeylerinin yorumlanması buna göre yapılmaktadır. Başarısızlık süresini etkileyen faktörleri belirlemek için Tablo 2 incelendiğinde, sigara tüketimi değişkeninin 4. düzeyinin, boyut değişkeninin 4. düzeyinin ve patolojik evre değişkeninin 3. ile 4. düzeylerinin %95 güven düzeyinde önemli olduğu görülmüştür ($p<0.05$).

Parametrik regresyon modellerini yorumlarken hızlandırma faktörünün yaşam süresinin azalmasını hızlandırıp hızlandırmadığı dikkate alınmaktadır. Buna göre Tablo 2.'deki bilgilerden yararlanarak sigara tüketimi 61'den fazla olan hastalara ait hızlandırma faktörü yaklaşık 0.35 olup, sigara tüketimi 61'den fazla olan hastaların yaşam süresi sigara tüketimi 5'ten az olan hastaların yaşam süresinden 2.9 kat daha kısadır biçiminde yorumlanabilir. Benzer biçimde tümör boyutu 50 mm'den büyük olan hastaların ortalama yaşam süresi tümör boyutu 30 mm'den küçük olan hastalara göre yaklaşık 2 kat, patolojik evresi Evre-IV olan hastaların ortalama yaşam süresi patolojik evresi Evre-I olan hastalara göre yaklaşık 4.6 kat daha kısadır.

3.3. Paylaşılmamış Zayıflık Modeli Sonuçları

Yaşam çözümlemesinde gözlenemeyen nedenlerden kaynaklanan heterojenliğin incelenmesinde zayıflık modelleri kullanılmaktadır. Bu çalışmada zayıflık teriminin paylaşılmamış ve paylaşılmış olduğu durumlar incelenmiştir. Paylaşılmamış zayıflık modelleri incelenirken en uygun parametrik regresyon modeli olarak belirlenen log- lojistik dağılım, Gamma ve Ters-Gauss zayıflık terimleri ile incelenmiş olup bu incelemelere ilişkin sonuçlar sırasıyla Tablo 3 ve Tablo 4'te verilmiştir.

Tablo 3. Gamma Zayıflık Terimi İçeren Log-Lojistik HBS Modeli Çözümlemesi

Değişken ve düzeyleri		$\hat{\beta}$	**Hızlandırma Faktörü**	**p değeri**
Kesişim terimi		5.0954	-	< 0.0001
Yaş	1. <=39 2. 40-49 3. 50-59 4. 60-69 5. >=70	- -0.1674 -0.2928 -0.7039 -0.1135	- 0,846 0,746 0,495 0,893	- 0.159 0.479 0.783 0.718
Sigara Tüketimi	1. <=5 2. 6-30 3. 31-60 4. 5>=61	- -0.4679 -0.3169 -0.9642	- 0,626 0,728 0,381	- 0.301 0.453 0.035
Genişletilmiş Rezeksiyon	1. Yok 2. Var	- -0.06006	- 0,548	0,052
Boyut	1. <=30 2. 31-40 3. 41-50 4. >=50	- 0.3354 -0.4412 -0.7572	- 1,398 0,643 0,469	- 0.264 0.156 0.004
İnvazyon	1. Yok 2. Var	- -0.02770	- 0,973	- 0.921
Patolojik Evre	1. Evre I 2. Evre II 3. Evre III 4. Evre IV	- -0.1950 -0.6418 -1.3255	- 0,823 0,526 0,266	- 0.515 0.016 0.001
Ölçek (σ)		0.5322		
Zayıflık terimi (θ)		0.9616		0.061
-2LogL	435.1918			

Tablo 4. Ters-Gauss Zayıflık Terimi İçeren Log-Lojistik HBS Modeli Çözümlemesi

Değişken ve düzeyleri		$\hat{\beta}$	**Hızlandırma Faktörü**	**p değeri**
Kesişim terimi		4.9835	-	< 0.0001
Yaş	1. <=39 2. 40-49 3. 50-59 4. 60-69 5. >=70	- -0.1849 -0.6946 -0.1173 -0.3026	- 0.831 0.499 0.890 0.739	- 0.690 0.471 0.779 0.164
Sigara Tüketimi	1. <=5 2. 6-30 3. 31-60 4. 5>=61	- -0.4651 -0.9474 -0.3127	- 0.628 0.388 0.731	- 0.303 0.459 0.041
Genişletilmiş Rezeksiyon	1. Yok 2. Var	- -0.6297	- 0.533	- 0.053
Boyut	1. <=30 2. 31-40 3. 41-50 4. >=50	- 0.3308 -0.4462 -0.753	- 1.392 0.640 0.471	- 0.272 0.149 0.004
İnvazyon	1. Yok 2. Var	- -0.0518	- 0.950	- 0.851

Değişken ve düzeyleri		$\hat{\beta}$	**Hızlandırma Faktörü**	**p değeri**
Patolojik Evre	1. Evre I 2. Evre II 3. Evre III 4. Evre IV	- -0.1714 -0.6041 -1.3511	- 0.842 0.546 0.259	- 0.569 0.028 0.001
Ölçek (σ)		0.4954		
Zayıflık terimi (θ)		2.0351		0.055
-2LogL	435.0314			

Gamma zayıflık terimi ve Ters-Gauss zayıflık terimi ile kurulan modeller %10 anlamlılık düzeyinde değerlendirilirse bu modellere ilişkin zayıflık terimlerinin anlamlı olduğu görülmektedir. Bu da zayıflık terimi içeren modellerin kullanımının daha uygun olduğunu göstermektedir. Zayıflık terimlerinin yorumlanmasında zayıflık teriminin heterojenliğin bir ölçüsü olduğu bilgisi altında Pankratz, Andrade ve Thernau (2005) yaklaşımına göre zayıflık teriminin karekökünün üstel ifadesi $\exp(\sqrt{\theta})$ rastgele etkinin yani zayıflığın etkisini göstermektedir. Log-lojistik dağılıma sahip Gamma zayıflık terimi ile kurulan model için $\exp(\sqrt{\theta})$ = $\exp(\sqrt{.9616}) = 2.67$ olarak elde edilmiştir. Bu değer bir hastanın, çalışma sürecinde seçilen bir noktada, tüm çalışma grubuna ait riskten 2.67 kata kadar daha fazla ya da az risk taşıyabileceğini ifade etmektedir. Benzer biçimde log-lojistik dağılıma sahip ters-Gauss zayıflık terimi ile kurulan model için $\exp(\sqrt{\theta})$ = $\exp(\sqrt{2.0351}) = 4.16$ olarak elde edilmiştir. Bu değer bir hastanın, çalışma sürecinde seçilen bir noktada, tüm çalışma grubuna ait riskten 4.16 kata kadar daha fazla ya da az risk taşıyabileceğini ifade etmektedir. Bu iki model karşılaştırılıken zayıflık teriminin anlamlılığı ve -2LogL beraber dikkate alınırsa, log-lojistik dağılıma sahip ters-Gauss zayıflık terimi ile kurulan modelin daha iyi olduğu söylenebilmektedir.

3.4. Paylaşılmış Zayıflık Modeli Sonuçları

Açıklayıcı değişkenlerin aynı düzeylere sahip bireyler bir grup olarak ele alınmış ve aynı zayıflık terimini paylaşacakları düşünülerek paylaşılmış zayıflık modelleri incelenmiştir. Bu modellerin, yaş, sigara tüketimi, genişletilmiş rezeksiyon, invazyon değişkenleri için anlamlı olmadığı görülmüştür. Boyut değişkeni ve patolojik evre için ise paylaşılmış zayıflık modelleri anlamlı bulunmuştur ($p < 0.05$). Bu değişkenlere ilişkin bilgiler Tablo 5 ve Tablo 6'da verilmiştir.

Tablo 5. Boyut Değişkeni için Zayıflık Terimi İçeren Log-Lojistik Hbs Modeli Çözümlemesi

Açıklayıcı Değişken		**Boyut**					
		Gamma			Ters Gauss		
		Parametre tahmini	Hızlandırma Faktörü	p-değeri	Parametre tahmini	Hızlandırma Faktörü	p-değeri
Yaş	1. <=39 2. 40-49 3. 50-59 4. 60-69 5. >=70	 0.1294 -0.1575 0.0030 -0.2071	 1.138 0.854 1.003 0.813	 0.777 0.708 0.995 0.675	 0.1274 -0.1569 0.0037 -0.2080	 1.136 0.855 1.003 0.812	 0.781 0.709 0.993 0.674
Sigara Tüketimi	1. <=5 2. 6-30 3. 31-60 4. >=61	 -0.6108 -0.4620 -1.2160	 0.543 0.630 0.300	 0.012 0.314 0.204	 -0.6062 -0.4559 -1.2126	 0.545 0.634 0.297	 0.208 0.321 0.012
Genişletilmiş Rezeksiyon	1. Yok 2. Var	 -0.2888	 0.750	 0.285	 -0.2920	 0.747	 0.283
İnvazyon	1. Yok 2. Var	 -0.3060	 0.736	 0.229	 -0.3018	 0.739	 0.236
Patolojik Evre	1. Evre I 2. Evre II 3. Evre III 4. Evre IV	 -0.1822 -0.6045 -1.5297	 0.833 0.546 0.217	 0.536 0.023 0.000	 -0.1811 -0.6034 -1.5330	 0.834 0.547 0.216	 0.538 0.023 0.000
Theta (Θ)		0.1152		0.027	0.1307		0.027
-2LogL		447.891			447.876		

Tablo 6. Patolojik Evre Değişkeni için Zayıflık Terimi İçeren Log-Lojistik Hbs Modeli Çözümlemesi

Açıklayıcı Değişken		**Patolojik Evre**					
		Gamma			Ters Gauss		
		Parametre tahmini	Hızlandırma Faktörü	p-değeri	Parametre tahmini	Hızlandırma Faktörü	p-değeri
Yaş	1. <=39						
	2. 40-49	-0.0416	0.959	0.930	-0.2151	0.806	0.664
	3. 50-59	-0.3232	0.724	0.444	-0.4052	0.667	0.367
	4. 60-69	-0.0886	0.915	0.835	-0.1272	0.881	0.768
	5. >=70	-0.4773	0.620	0.327	-0.5904	0.554	0.236
Sigara Tüketimi	1. <=5						
	2. 6-30	-0.4514	0.637	0.323	-0.3736	0.688	0.404
	3. 31-60	-0.4083	0.665	0.340	-0.3636	0.695	0.380
	4. 5>=61	-1.027	0.358	0.029	-0.9470	0.388	0.043
Genişletilmiş Rezeksiyon	1. Yok						
	2. Var	-0.3607	0.697	0.264	-0.6004	0.549	0.191
Boyut	1. <=30						
	2. 31-40	0.4217	1.525	0.165	0.4812	1.618	0.118
	3. 41-50	-0.4182	0.658	0.181	-0.3998	0.670	0.186
	4. >=50	-0.6330	0.531	0.021	-0.6380	0.528	0.021
İnvazyon	1. Yok						
	2. Var	-0.0865	0.917	0.747	-0.0044	0.996	0.987
Theta (Θ)		0.514		<0.0001	2.104		<0.0001
-2LogL		446.972			445.441		

Tablo 5'de boyut değişkeni bir küme olarak değerlendirilmiş olup kitle ikiye ayrılmıştır, boyut değişkeni için referans grubu olarak boyut 4 (>=50) belirlenmiştir. Kümedeki tüm bireylerin aynı zayıflığı paylaştığı varsayımı altında paylaşılmış zayıflık modeli kurulmuş ve model 0.05 anlamlılık düzeyinde önemli bulunmuştur. Gamma ve ters Gauss zayıflık terimlerini içeren her iki model için de zayıflık teriminin etkisi yaklaşık olarak $\exp(\sqrt{\theta}) = 1.4$ elde edilmiştir. Buna göre tümörün boyutu >=50 olan bireylerin çalışma grubundaki diğer bireylere göre 1.4 kata kadar daha fazla ya da daha az risk taşıdığı söylenebilmektedir.

Benzer biçimde; Tablo 6'da patolojik evre değişkeni bir küme olarak değerlendirilmiş olup kitle, patolojik evre değişkeni için referans grubu Evre-IV olmak üzere ikiye ayrılmıştır. Paylaşılmış zayıflık modeli, zayıflık terimi Gamma ve ters Gauss olmak üzere kurulmuş ve model istatistiksel açıdan anlamlı bulunmuştur. Zayıflık terimi Gamma olan paylaşılmış zayıflık modeli için yaklaşık olarak $\exp(\sqrt{\theta}) = 2$ elde edilmiştir. Bu değer patolojik evresi Evre-IV olan bireylerin çalışma grubundaki bireylere göre 2 kata kadar daha fazla ya da daha az risk taşıdığını ifade etmektedir. Bir diğer model olan, ters Gauss zayıflık terimli paylaşılmış zayıflık modelinde ise $\exp(\sqrt{\theta}) = 4.25$ olarak elde edilmiştir. Bu model dikkate alınarak, patolojik evresi Evre-IV olan bireylerin çalışma grubundaki diğer bireylere göre 4.25 kata kadar daha fazla ya da daha az risk taşıdığını ifade etmektedir.

4. SONUÇ

Yaşam çözümlemesinde gözlenemeyen değişkenlerden kaynaklanan heterojenliğin açıklaması zayıflık modelleri ile yapılabilmektedir. Yaşam çözümlemesi kullanılarak elde edilen bulguların tutarlılığı için heterojenliğin bireylere ya da gruplara etkisinin incelenmesi gerekmektedir. Heterojenliğin yüksek olduğu durumlarda çalışma grubuna dair çözümleme yapılırken zayıflık modelleri göz önünde bulundurulmalıdır.

Akciğer kanseri tedavisi gören 236 bireye ait veri kümesine yaşam çözümlemesi yöntemleri uygulanmıştır. Zayıflık terimi içeren modeller ayrıntılı olarak incelenmiştir. Yaşam süresinin log-lojistik dağılıma uyum sağladığı gözlenmiş olup çözümlemeler bu dağılım üzerinden çeşitlendirilmiştir. Paylaşılmamış zayıflık modeli ve log-lojistik HBS modeli karşılaştırıldığında, paylaşılmamış zayıflık modelinin veriyi yorumlamada daha iyi olduğu görülmüştür. Paylaşılmış zayıflık modelleri ise verinin açıklayıcı değişkenlere göre gruplandırılması durumunda grupların birbirinden ne oranda az ya da fazla risk taşıdığını görebilmek ve yorumlayabilmek için kullanılmıştır.

Ülkemizde, özellikle kanser verileri yaşam çözümlemesi yaklaşımı ile incelenirken verinin homojen olduğu varsayımı yapılarak orantılı tehlikeler modelinden yararlanılmaktadır. Ancak, gerçekte incelenen verinin homojen olmadığı durumlarla da karşılaşılmaktadır. Yaşam verisinin homojen olmaması, çözümleme sonucunda ulaşılan değerlendirmelerin etkinliğinin azalmasına ve verinin yeterince açıklanamamasına neden olmaktadır. Bu nedenle, yaşam çözümlemesi verilerinde heterojenliğinin açıklanabilmesi için zayıflık modellerinin göz önüne alınması gerekmektedir.

References

Ata, N. (2005). Yaşam Çözümlemesinde Orantısız Hazard Modeli, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, Ankara, 2005.

Babiker A., Cuzick J. (1994). A simple frailty model for family studies with covariates. Statistics in Medicine, 13, 1679-1692.

Clayton, D. (1978). A model for association in bivariate life tables and ıts applications in epidemiological studies of familial tendency in chronic disease ıncidence. Biometrika, 65, 141–151.

Clayton, D., Cuzick, J. (1995). Multivariate generalisations of the proportional hazards model (with discussion). Journal of the Royal Statistical Society, Series A, 148, 82–117.

Congdon, P. (1995). Modelling Frailty İn Area Mortality. Statistics in Medicine, 14, 1859-1874.

Cox, D.R. (1972). Regression models and life-tables. Journal of the Royal Statistical Society,Series B, 34, 187-220.

Duchateau, L., Janssen, P. (2007). The Frailty Model, Springer, New York.

Economou, P., Caroni, C. (2005). Graphical tests for the assumption of gamma and inverse Gaussian frailty distributions. Lifetime Data Analysis 11, 565–582.

Guo, G., Rodriguez, G. (1992). Estiamting amultivariate proportional hazards model for clustered data using the EM algorithm with an application to child survival in Guatemala. Journal of the American Statistical Association, 87, 969-976.

Gutierrez, R.G. (2002). Parametric frailty and shared frailty survival models. The Stata Journal, 2, 22–44.

Hougaard, P.(1984). Life table methods for heterogeneous populations. Biometrika, 71, 75–83.

Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. Biometrika 73, 387–396.

Hougaard, P.(1995). Frailty models for survival data. Lifetime Data Analysis, 1, 255–273.

Hougaard, P. (2000). Analysis of Multivariate Survival Data, Springer, New York.

Ibrahim, J. G., Chen, M., Sinha, D. (2001), Bayesian Survival Analysis, Springer, New York.

Keiding, N., Andersen, P., Klein, J. (1997). The role of frailty models and accelerated failure time models in describing heterogeneity due to omitted covariates. Statistics in Medicine 16, 215–224.

Kheiri, S., Kimber, A., Meshkani, M.R. (2007). Bayesian analysis of an inverse Gaussian correlated frailty model, Computational Statistics and Data Analysis 51, 5317–5326.

Klein, J.P. (1992). Semiparametric estimation of random effects using the Cox model based on the EM algorithm. Biometrics 48, 795–806.

Klein, J. P. and Moeschberger, M. L. (2003). Survival Analysis: Techniques for Censored and Truncated Data, 2nd edition, Springer, New York.

Lanchester, T. (1979). Econometric methods for the duration of unemployment. Econometrica, 47, 939–956.

McGilchrist, C.A., Aisbett, C.W. (1991). Regression with frailty in survival analysis. Biometrics 47, 461–466.

Oakes, D. (1982). A concordance test for independence in the presence of censoring. Biometrics 38, 451–455.

O'Quigley J,, Stare, J. (2002). Proportional hazards models with frailties and random effects. Statistics in Medicine, 21, 3219-3233.

O'Quigley J,, Stare, J. (2004). Fit and frailties in proportional hazards regression. Statistics in Medicine, 21, 3219-3233.

Pankratz, V. S., de Andrade, M. and Thernau, T. M. (2005). Random effects Cox proportional hazard model: general variance components methods for time-to-event data. Genetic Epidemology, 28, 97-109.

Price, D.L., Manatunga, A.K. (2001). Modelling survival data with a cured fraction using frailty models. Statistics in Medicine 20, 1515–1527.

Sahu, S.K., Dey, D.K., Aslanidou, H., Sinha, D. (1997). A weibull regression model with gamma frailties for multivariate survival data. Lifetime Data Analysis, 3, 123–137.

Vaupel, J.W., Manton, K., Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. Demography, 16, 439–454, 1979.

Whitmore, G.A, Lee, M.L.T. (1991). A multivariate survival distribution generated by an ınverse gaussian mixture of exponentials, Technometrics, 33, 39–50.

Wienke A. (2011). Frailty Models in Survival Analysis, Chapman&Hall, Florida.

Wienke, A., Christensen, K., Holm, N., Yashin, A. (2000). Heritability of death from respiratory diseases: an analysis of Danish twin survival data using a correlated frailty model. In: Medical Infobahn for Europe. A. Hasman et al. (eds.), IOS Press, Amsterdam, 407–411.

Wienke, A., Yashin, I.A., Zdravkovic, S., Pedersen, N.L., Marenberg, M.E., De Faire, U. (2004). Genetic influences on CHD-Death and the impact of known risk factors: Comparison of two frailty models. Behavior Genetics, 34, 585-592.

Xue, X., Brookmeyer, R. (1996). Bivariate frailty model for the analysis of multivariate survival time. Lifetime Data Analysis 2, 277-289.

Yashin, A.I., Iachine, I.A. (1995). Genetic analysis of durations: correlated frailty model applied to survival of danish twins. Genetic Epidemiology, 12, 529 – 538.

Yashin, A.I., Vaupel, J.W., Iachine, I.A. (1995). Correlated individual frailty: An advantageous approach to survival analysis of bivariate data. Mathematical Population Studies, 5, 145 – 159.

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2016.04.01.OR.04

DETERMINATION OF EFFECTIVE CRITICAL FACTORS IN SUCCESSFUL EFFICIENCY MEASUREMENT OF UNIVERSITY DEPARTMENTS BY USING FUZZY DEMATEL METHOD

Turgay TÜRKER*

Department of Industrial Engineering, Karabük University, Karabük, Turkey

Murat ETÖZ †

Department of Business Administration, Mediterrian University, Antalya, Turkey

Yasemin ALTUN TÜRKER‡

Department of Industrial Engineering, Karabük University, Karabük, Turkey

Received: 22 February 2016

Accepted: 26 May 2016

Abstract

Successful efficiency measurement of university departments is very important issue in today's globalised world and depends on paying high attention on critical input and output factors affecting efficiency measurement. In this study we aim to determine critical input and output factors in efficiency measurement of university departments by using fuzzy DEMATEL (Decision Making Trial and Evaluation Laboratory) method. Factors that are used in efficiency measurement of university departments have been extracted from the literature. Then Fuzzy DEMATEL method has been employed to separate cause and effect group of factors in input and output sets. Cause groups are advised to be used in DEA (Data Envelopment Analysis). This study is the first reference in the literature which uses a fuzzy DEMATEL technique in determination of effective critical input and output factors in successful efficiency measurement of university departments.

Keywords: Efficiency measurement, Data envelopment analysis, Fuzzy DEMATEL

ÜNİVERSİTELERDE BÖLÜM ETKİNLİKLERİNİN ÖLÇÜLMESİNDE KULLANILAN KRİTİK BAŞARI FAKTÖRLERİNİN BULANIK DEMATEL YÖNTEMİ İLE BELİRLENMESİ

Özet

Günümüzün globalleşen dünyasında üniversite bölümlerinin başarılı etkinlik ölçümü çok önemli bir konudur ve bu ölçümün başarısı doğru girdi ve çıktı faktörlerinin belirlenmesine bağlıdır. Bu çalışmada, üniversite bölümlerinin etkinlik ölçümünde kullanılan kritik girdi ve çıktı faktörleri Bulanık DEMATEL (Decision Making Trial and Evaluation Laboratory) yöntemi kullanılarak belirlenmiştir. Öncelikle literatürden üniversite bölümlerinin etkinlik ölçümünde kullanılan faktörler elde edilmiştir. Daha sonra Bulanık DEMATEL yöntemi ile girdi ve çıktı kümeleri için ayrı ayrı etkileyen-etkilenen gruplar belirlenmiştir. Etkileyen grupların VZA'da (Veri Zarflama Analizi) kullanılması önerilmiştir. Bu çalışma üniversite bölümlerinin etkinliklerinin ölçülmesinde kullanılan kritik girdi ve çıktı faktörlerinin tayini için bulanık DEMATEL tekniğinin kullanıldığı literatürdeki ilk çalışmadır.

Anahtar Kelimeler: Etkinlik ölçümü, Veri zarflama analizi, Bulanık DEMATEL

* tturker@karabuk.edu.tr (Corresponding author)

† muratetoz@yahoo.co.uk

‡ yaseminaltun@karabuk.edu.tr

1. INTRODUCTION

Economic crises caused by increase in public sector spending has led to a questioning the concept of public financial management in many countries. Improving the quality of public services, increasing the resource usage capacity, ensuring efficiency, effectiveness and thriftiness in resource usage, make nonprofit government organizations employing efficiency analysis methods in their operations (Eruz, 2005).

Higher education is one of the most important factors in producing information, in high quality servicing to the community and in meeting the qualified work force demand of a country. Transition process to information society began at the last quarter of the twentieth century throughout the world. This process has been continuing since then. Transition process has brought a new global economic structure called knowledge economy. In this new structure, while economic power of individuals is measured with knowledge, education levels and competitiveness of countries is measured with the human and social capital. This process has increased the expectations of public from universities which are primarily responsible for knowledge generation and sharing and has led to increased competition between universities (YÖK, 2007). Besides expectations, the demand for higher education has been continuously increasing in developing countries which have high percentage of young population. While this situation increases the number of universities in developing countries, universities are forced to use their resources effectively. For example, in Turkey, number of universities was 76 in 2003. This number has increased to 190 in 2015. Similarly while enrolments in higher education were about 1,200,000 in 2003, this number reached to about 6,000,000 in 2015. Therefore analysis of effectiveness and efficiency in universities has become a major management tool.

Quality of education is prerequisite for the continuity of competitive advantage of universities. Sustainability of quality depends on efficiency measurements. Efficiency of universities and university departments has been researched from many different points of view by many researches (Ahn, Arnold, Charnes & Cooper, 1989; Arcelus & Coleman, 1997; Athanassopoulos & Shale, 1997; Avkiran, 2001; Breu & Raab, 1994; Coelli, 1996; Coelli, Rao & Battese, 1998; El Mahgary & Lahdelma, 1995; Flegg, Allen, Field & Thurlow, 2004; Flegg & Allen, 2007; Friedman & Sinuany-Stern, 1997; Haksever & Muragishi,1998; Johnes, 2006; Johnes & Yu, 2008; Koksal & Nalcacı, 2006; McMillan & Datta, 1998; Preeti, Shiv & Singh, 2009; Worthington & Lee, 2008). These researches basically strive to get answers to the following questions: Do the inputs used in the system maximize the resulting outputs? Are the types and quantities of inputs and their relevant outputs appropriate to achieve educational goals of universities or university departments? Under current budgetary constraints are academic objectives complied? If the budget cuts are made, to what extent the quality of education to be lowered is allowed? Finding appropriate answers to these questions is the core subject attracted the attention of many researchers. It is clear that academic units' efficiency evaluation is a difficult task due to the difficulties in measurements of inputs and outputs, to the large number of such measures relative to the departments, to the data collection problems, and to the implications of resource allocations as a result of budgetary cuts which are difficult to estimate in advance. Comparing different university departments in terms of efficiency make this problem even more complicated (Arcelus & Coleman, 1997).

Efficiency is mostly measured by using parametric methods (i.e. ordinary least square method), nonparametric methods (i.e. Data Envelopment Analysis) and performance indicators. These methods have pros and cons comparing to each other. Performance indicators are an effective method when only one input and one output are determined in measuring efficiency. In case of multi input and output, the method doesn't work. In the literature, some of the researches on university efficiency measurement that focus on performance indicators are Cave, Hanney & Kogan, 1991; Chalmers, 2008; OECD, 2007; Pereira & Tavares, 2002; Ward, D. 2007. Parametric methods need distribution functions for technology and inefficiency. On the other hand, non-parametric methods use multiple inputs and outputs in efficiency evaluation.

DEA is a nonparametric method which is widely used in measuring the efficiency of organizations. Comparing to other measurement methods, it's advantages is lied on the fact that it uses distance function approach in which multiple inputs and outputs can be employed (Johnes & Yu, 2008). There are considerable amount of studies in the literature which use DEA method in efficiency evaluation of universities and university departments (Athanassopoulos & Shale, 1997; Avkiran, 2001; Coelli, 1996; Flegg, Allen, Field & Thurlow, 2004; McMillan & Datta, 1998; Casu & Thanassoulis, 2006; Fox & Milbourne, 1999; Salerno, 2006; Stevens, 2001; Worthington, 2001; Glass, McKillop & Hyndman, 1995; Johnes, 1996).

DEA is a linear programming based technique that aim to evaluate relative performances of decision making units (DMU) in case there exists multiple inputs and outputs which have different measurement units. The

Basic assumptions of DEA are that all DMUs have similar strategic objectives and they use same kind of inputs and produce same kind of outputs. DEA measures efficiency of DMUs that have same goals and objectives. By using DEA, inefficient DMUs can be determined and corrective measures can be taken. Evaluations of many input and output at the same time are very difficult task for managers. At this point, DEA provides important tool in determination of relative efficiency (Oruc, 2008).

The most important part of DEA is the selection and definition of proper input and output factors. In the literature mostly used inputs in university departments' efficiency evaluation are number of academic staff, number of assist academic staff, number of non-academic staff, number of undergraduate students, number of MSc. students, number of PhD. students, total department expenditures, research budget, department budget, number of laboratories, number of computers, number of classrooms, building usage area, number of courses and minimum score of student selection exam. Outputs are number of papers published in academic journals (SCI, SSCI), number of papers published in peer-reviewed journals, number of proceedings, number of academic researches, number of undergraduate awards, number of postgraduate awards, number of doctorates awards and research Incomes. Numerous numbers of input-output combinations were formed in previous studies (Ahn, Arnold, Charnes & Cooper, 1989; Arcelus & Coleman, 1997; Athanassopoulos & Shale, 1997; Avkiran, 2001; Breu & Raab, 1994; Flegg, Allen, Field & Thurlow, 2004; Friedman & Sinuany-Stern, 1997; Johnes, 2006; Johnes & Yu, 2008; Koksal & Nalcacı, 2006; McMillan & Datta, 1998; Preeti, Shiv & Singh, 2009; Worthington & Lee, 2008, Beasley,1990,1995; Bessent, Bessent, Charnes, Cooper & Thorogood, 1983). Moreover, data availability is the other major concern in formation of input-output factors (Katharaki & Katharakis, 2010).

However, a systematic method to identify the inputs and outputs hasn't been observed in the literature. Instead, researchers determine these factors intuitively or conceptually. The major drawback in selection of these factors is that researches ignore effects of factors to each other. Factors that are influenced from other factors are not useful in measuring the efficiency of DMUs. Therefore critical factors in efficiency measurement must be determined. Discrimination of factors as cause and effect group is the subject of DEMATEL method. Gabus and Fontela developed DEMATEL method (Gabus & Fontela, 1972, 1973). This method visualizes the complicated causal relationships of factors via matrices and digraphs. DEMATEL also separates the criteria into cause and effect factors' groups which make decision making easier (Wu & Lee, 2007).

It is usually difficult to make decision with crisp values in vague environment. Uncertainty in decision making necessitates employment of fuzzy logic. Fuzzy logic usually used to handle uncertainty and vagueness of decision making process (Zadeh, 1965; Bellman & Zadeh, 1970). DEMATEL method combined with Fuzzy logic is used to obtain more realistic decisions in fuzzy environments. In the literature, there exist some studies that combine DEMATEL and Fuzzy logic (Wu & Lee, 2007; Mokhtarian,2007; Zhou, Huang & Zhang, 2011; Chang, Chang & Wu, 2011).

In this study we aimed to determine critical input and output factors by using fuzzy DEMATEL method for efficiency measurement of university departments. Factors that are used in efficiency measurement of university departments were extracted from the literature. Then Fuzzy DEMATEL method was employed to separate cause and effect group of factors in input and output sets. Cause groups were advised to be used in DEA analysis.

The following sections of the paper are organized as follows. Literature Review about university departments' efficiency measurements and applications of DEMATEL-fuzzy logic combination are presented in section 2. In section 3, applied method is described. In section 4, empirical study is presented. Section 5 is dedicated to discussion about results of the study. Conclusion and further study opportunities are presented in the last section.

2. LITERATURE REVIEW

In recent years, considerable amount studies have been done to analyze the efficiency of academic departments in universities. Among these, some important studies are briefly reviewed as follows.

Bessent et al. (1983) used DEA to evaluate the relative efficiency of education programs in a community college in USA. Outputs were considered as revenue from state government, number of students completing the program, and employer satisfaction with training of students. Inputs were student contact hours, number of full time equivalent instructors, square feet of facilities for each program and direct instructional expenditure.

Beasley (1990) studied efficiency of Chemistry and Physics departments in UK. Financial variables (research income and expenditure) were used as inputs. Outputs were numbers of undergraduate and postgraduate students and research ratings. Beasley (1995) also used same data set to determine the research and teaching efficiencies where weight restrictions were applied.

Arcelus and Coleman (1997) assessed short and long term effects of fixed budget formula on the efficiency of academic departments of New Brunswick University by using DEA. Inputs of their model were the number of full time equivalents teaching staffs (professors, instructors etc.), number of support staffs (secretaries, non-faculty student advisors, technicians etc.), operating expenses and library expenses. The corresponding outputs were average enrollment per class (student/class), average number of classes taught per department (sections), number of undergraduate students and number of undergraduate students receiving their degree in a given year at any given department, number of graduate students in each departments program receiving their degree in a given year (MSc. and PhD.)

Köksal and Nalçacı (2006) used DEA to measure relative efficiencies of academic departments of an engineering college. They determined input and output criteria utilizing the academic personnel performance measurement of the collage. They also developed measures to compare departments of different disciplines. Their input measures were academic staff salaries (Sum of salaries of faculty and assistants in the department in the period under consideration), potential of the department (20x(average number of publications of a faculty member when appointed as an associate professor, over the past 10 years in the department), Entering students (Weighted sum of students starting BSc. education four years before (1 point), M.Sc. education three years before (3 points) and PhD. education five years before (6 points) than the measurement period).Output measures were research activities and quality (sum of points gathered by the departments academic personnel in that period in publications, editorial work and translation), education activities and quality (Sum of points gathered by the department personnel in that period due to their education activities and their quality), other activities (activities such as administrative duties in the university or other institutions, conference organizations and seminars) and Graduates (Weighted sum of BSc. (1 point), MSc. (3 points) and PhD. (6 points) students graduated in a period).

Preeti et al. (2009) evaluated the performance efficiencies of 19 academic departments of IIT Roorkee (India) through DEA technique. Outputs that they used were the number of academic staff, the number of Non-academic staff and departmental operating cost. Outputs were total enrolled students, number of postgraduate degrees and number of graduate degrees and research index.

After Gabus and Fontela (1972, 1973) initiated DEMATEL technique, considerable amount of studies have been done concentrated on structural model that gathers group knowledge and visualize the causal relationship of criteria by using graphical diagram. Recently, the DEMATEL and fuzzy DEMATEL methods have been studied in different areas, such as real estate agent service quality (Tseng, 2009), personnel selection (Aksakal & Dağdeviren, 2010), emergency management (Zhou, Huang, & Zhang, 2011), supplier selection and evaluation (Chang, Chang, & Wu, 2011; Büyüközkan & Çifçi, 2012; Mavi & Shahabi, 2015), human resources management (Chou, Sun, & Yen, 2012), machine selection (Organ, 2013), facility layout (Altuntas, Selim, & Dereli, 2014), private school evaluation (Baykasoglu & Durmusoglu, 2014), determining environmental performance (Tsai et al., 2015), project selection (Ortíz, Felizzola, & Isaza, 2015; Vinodh & Swarnakar, 2015), ship selection (Sener, 2016).

Aksakal & Dağdeviren (2010) proposed an integrated algorithm by combined DEMATEL and Analytic Network Process (ANP) for decision making about the personnel selection. There were six evaluation criteria for four candidate personnel. The evaluation criteria were experience, communication, foreign language, computer skills, team member skills, strategic thinking. The analyze results indicated that, communication, team member skills and experience are the crucial criteria in personnel selection, respectively.

Zhou et al. (2011) used fuzzy DEMATEL method to determine influencing critical success factors of emergency management. In this study, a set of 20 complex influencing factors were divided into a cause group and an effect group. According to the analyze results, five factors were defined critical success factors in emergency management.

Chang et al. (2011) studied fuzzy DEMATEL method in determining key factors in supplier selection. Their model analyzed supplier performance to determine influencing critical success factors in supplier selection. In the study, a fuzzy DEMATEL questionnaire was sent to seventeen professional purchasing personnel in the electronic industry. The results showed that technology ability, stable delivery of goods, lead-time and production capability criteria are more influential than other evaluation criteria.

Chou et al. (2012) used an integration of fuzzy Analytic Hierarchy Process (AHP) and fuzzy DEMATEL method in human resource for science and technology (HRST). The fuzzy DEMATEL method was used to capture the complex relationships between dimensions and criteria. It was seen that for HRST, infrastructure is the most important criteria and the education, R&D, expenses and immediate output are more important second-tier criteria than value, cooperation, labor market, human capital and intermediate output.

Organ (2013) used fuzzy DEMATEL method to reveal the relationships among each criteria in machine selection problem for a textile firm. The used criteria were cost, quality, technical properties, performance and flexibility of machines. It was found that, the most important criteria for selection machine in firms is technical properties.

Altuntas et al. (2014) presented a fuzzy DEMATEL-based solution approach for facility layout problem. The proposed approach allowed both qualitative and quantitative location factors. They considered six important location factors. These are material flow, information flow, personnel flow, equipment flow, environmental condition and supervision of personnel. To explore the viability of the proposed approach, a real world problem in a machinery industry firm was handled.

Baykasoglu & Durmusoglu (2014) used DEMATEL based on ANP and Fuzzy Cognitive Map model for private primary school selection problem. The DEMATEL method was used to determine the relative weight of the main criteria. A case study in Turkey was handled.

Tsai et al. (2015) used to fuzzy DEMATEL to define the direction and level of interaction between environmental performance criteria of the Printed Circuit Board (PCB) industry in Taiwan. The study classified the environmental performance measurement system into 4 major dimensions (i.e., green development, green manufacturing, green management and green recycling), which comprised 10 criteria (i.e., green design, green material procurement, air & water pollution, waste pollution, energy consumption, green marketing, green transport, green image, green packaging and product recycling). According to the results, it emerged that green design, green material procurement, and energy consumption are the most crucial criteria.

Ortiz et al. (2015) presented both an integrated DEMATEL and ANP technique and only ANP technique to establish the most suitable six sigma project in a public medical center. The hierarchical evaluation model included 3 strategies, 4 criteria and 15 sub-criteria and after analyzing, the integrated method and ANP were compared to evaluate their performance in the decision making process. Because of decreasing error probability, the integrated method could be found better.

Sener (2016) used DEMATEL method to determine ship selection criteria in maritime transportation industry. Real-world data was used to illustrate the application of the model. In this study, ten criteria were used. These were cost, payment due date, delivery time, reputation of the shipping company, flag of the ship, age of the ship, duration of detentions, classification organization, capacity of the ship and speed of the ship. According to the results, reputation of the company was the most important criteria and other critical factors were the duration of detentions, classification organization, cost, and age of the ship.

Although fuzzy DEMATEL method was employed in determination of critical factors in several fields, similar applications weren't encountered on the scope of university department's efficiency measurements. This study aimed to fill this gap in the literature.

3. Fuzzy DEMATEL Method

Fuzzy DEMATEL technique includes 7 basic steps;

***Step 1*:** Objectives and Evaluation criteria with respect to them are determined.

***Step 2*:** Decision makers are questioned to determine their judgments about the relationship between criteria. Since human judgments on evaluation criteria include uncertainty, five linguistic terms "Very high influence, High influence, Low influence, Very low influence, No influence" are determined. Then these linguistic terms are expressed as positive triangular fuzzy numbers as shown in Table 1. The answers of decision makers in terms of linguistic terms are converted to triangular fuzzy numbers.

Table 1. Fuzzy Linguistic Scale

Linguistic term	**Triangular fuzzy number**
No Influence (No)	(0, 0, 0.25)
Very Low Influence (VL)	(0, 0.25, 0.50)
Low Influence (L)	(0.25, 0.50, 0.75)
High Influence (H)	(0.5, 0.75, 1.0)
Very High Influence (VH)	(0.75, 1.0, 1.0)

Step 3: Let $\tilde{X}^{(k)}$ is the k. evaluators' fuzzy decision matrix about the criteria expressed in terms of fuzzy triangular numbers. $\tilde{X}^{(k)}$ is normalized as follows

$$\tilde{X}^{(k)} = \begin{bmatrix} 0 & \tilde{x}_{12}^{(k)} & \cdots & \tilde{x}_{1n}^{(k)} \\ \tilde{x}_{21}^{(k)} & 0 & \cdots & \tilde{x}_{2n}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{n1}^{(k)} & \tilde{x}_{n2}^{(k)} & \cdots & 0 \end{bmatrix}, \quad k = 1, 2, ..., p \tag{1}$$

$$r^{(k)} = \max_{1 \le i \le n} \left(\sum_{j=1}^{n} u_{ij}^{k} \right) \tag{2}$$

$$\tilde{x}_{ij}^{(k)} = \frac{\tilde{z}_{ij}^{(k)}}{r^{(k)}} = \left(\frac{l_{ij}^{(k)}}{r^{(k)}}, \frac{m_{ij}^{(k)}}{r^{(k)}}, \frac{u_{ij}^{(k)}}{r^{(k)}} \right) \tag{3}$$

Step 4: In this step, average value of p evaluators' normalized fuzzy decision matrix is found.

$$\tilde{X} = \frac{\tilde{X}^{(1)} \oplus \tilde{X}^{(2)} \oplus ... \oplus \tilde{X}^{(p)}}{p} \tag{4}$$

$$\tilde{X} = \begin{bmatrix} \tilde{x}_{11} & \tilde{x}_{12} & \cdots & \tilde{x}_{1n} \\ \tilde{x}_{21} & \tilde{x}_{22} & \cdots & \tilde{x}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{x}_{n1} & \tilde{x}_{n2} & \cdots & \tilde{x}_{nn} \end{bmatrix}; \quad \tilde{x}_{ij} = \frac{\sum_{k=1}^{p} \tilde{x}_{ij}^{(k)}}{p} \tag{5}$$

Step 5: After finding initial direct relation matrix and normalizing it, Total relation fuzzy matrix ($\tilde{T}$) is defined as follows;

$$\tilde{T} = \tilde{X}.(I - \tilde{X})^{-1} \tag{6}$$

Step 6: In this step $\tilde{D}_i$ and $\tilde{R}_i$ are calculated. $\tilde{D}_i$ is the sum of the row and $\tilde{R}_i$ is the sum of the column of $\tilde{T}$. Then $\tilde{D}_i$ and $\tilde{R}_i$ are defuzzified separately.

Best Nonfuzzy Performance (BNP) value (Hsieh et al., 2004) was used as a defuzzification procedure. The BNP value can be found using the following equation;

$$BNP_{ij} = \left((U_{ij} - L_{ij}) + (M_{ij} - L_{ij}) \right) / 3 + L_{ij} \qquad (7)$$

BNP_{ij} represents the defuzzified value of $\tilde{D}_i$ and $\tilde{R}_i$. We call defuzzified value of $\tilde{D}_i$ and $\tilde{R}_i$ as D_i and R_i respectively.

In order to determine causal relationships between Critical success factors, $D_i + R_i$ and $D_i - R_i$ are calculated. While $D_i + R_i$ represents degree of central role (how much importance the criteria has), $D_i - R_i$ shows the degree of relation. Relation divide the criteria in to cause and effect group. If $D_i - R_i$ is positive then criteria belong to cause group. If $D_i - R_i$ is negative then criteria belong to effect group.

Step 7: Causal diagram is constructed. In this diagram the horizontal axis represents $D_i - R_i$ while vertical axis represents $D_i + R_i$. In this diagram, Criteria above the horizontal axis mean that they belong to cause group. Criteria below the horizontal axis mean that they belong to effect group.

4. Empirical Study

In our empirical study, critical input and output factors were distinguished from input and output sets that are used in efficiency measurement of university departments by employing fuzzy DEMATEL method. Factors that are used in efficiency measurement of university departments were extracted from the literature.

According to our objective of determining critical factors in input and output sets, 15 input factors and 8 output factors were selected. Selected inputs and outputs are shown in Table 2 and Table 3, respectively.

Table 2. Selected Input Factors

Factor	Description
C_{in}1	Number of Academic Staff
C_{in}2	Number of Assist Academic Staff
C_{in}3	Number of non-Academic Staff
C_{in}4	Number of Undergraduate Students
C_{in}5	Number of MSc. Students
C_{in}6	Number of PhD. Students
C_{in}7	Total Department Expenditures
C_{in}8	Research Budget
C_{in}9	Department Budget
C_{in}10	Number of Laboratories
C_{in}11	Number of Computers
C_{in}12	Number of Classrooms
C_{in}13	Number of Square Meters of Building
C_{in}14	Number of Course
C_{in}15	Minimum Score of Student Selection Exam

Table 3. Selected Output Factors

Factor	Description
C_{out}1	Number of Papers Published in Academic Journals (SCI, SSCI)
C_{out}2	Number of Papers Published in Peer-reviewed Journals
C_{out}3	Number of Proceedings
C_{out}4	Number of Academic Researches
C_{out}5	Number of Undergraduate Awards
C_{out}6	Number of Postgraduate Awards
C_{out}7	Number of Doctorates Awards
C_{out}8	Research Incomes

Then questionnaire form was designed for taking opinions of experts who took active roles in academic departments. Experts were chosen from different universities and from different departments as much as possible. Under these circumstances, 8 experts were asked to make pair wise comparisons between each pair of critical input and output factors separately. Experts identified their opinions through linguistic scale determined in Table 1. As an example, the assessment data of an expert for input and output factors are shown in Table 4 and Table 5.

Table 4. Assessment Data of an Expert for Input Factors in Linguistic Scale

	Cin1	Cin2	Cin3	Cin4	Cin5	Cin6	Cin7	Cin8	Cin9	Cin10	Cin11	Cin12	Cin13	Cin14	Cin15
Cin1	-	H	L	L	L	L	L	L	L	No	VL	No	VL	VH	VL
Cin2	No	-	L	VL	L	L	VL	VL	VL	No	No	No	VL	No	No
Cin3	No	No	-	No	No	No	L	No	L	No	No	No	L	No	No
Cin4	VH	VH	H	-	H	H	H	L	VH	VH	VH	VH	VH	H	L
Cin5	VH	H	H	No	-	VH	H	H	H	H	H	H	H	VH	No
Cin6	H	L	L	VL	VL	-	L	H	H	H	H	H	H	H	No
Cin7	H	H	No	H	H	H	-	H	H	H	H	H	H	No	VL
Cin8	H	H	No	No	H	H	VH	-	VH	H	H	VL	No	No	No
Cin9	H	H	H	H	H	H	VH	VH	-	VH	H	H	H	L	VH
Cin10	L	L	No	H	H	H	VH	VH	VH	-	VL	VL	H	H	H
Cin11	No	No	No	H	H	H	H	No	VH	L	-	L	H	H	L
Cin12	No	No	No	VL	VL	VL	L	No	L	No	No	-	H	No	No
Cin13	No	No	No	H	H	H	VL	No	L	VL	VL	L	-	No	No
Cin14	H	L	No	No	No	No	L	No	L	L	L	H	VL	-	L
Cin15	VL	VL	No	VL	No	No	No	No	No	No	No	No	No	No	-

Table 5. Assessment Data of an Expert for Output Factors in Linguistic Scale

	$C_{out}1$	$C_{out}2$	$C_{out}3$	$C_{out}4$	$C_{out}5$	$C_{out}6$	$C_{out}7$	$C_{out}8$
$C_{out}1$	-	VH	H	L	No	No	No	VL
$C_{out}2$	H	-	L	H	No	VL	No	No
$C_{out}3$	L	VL	-	No	No	No	No	No
$C_{out}4$	H	H	VH	-	VL	H	H	VH
$C_{out}5$	L	VL	L	L	-	VH	H	VL
$C_{out}6$	H	H	VH	H	No	-	VH	H
$C_{out}7$	H	VH	H	L	No	No	-	H
$C_{out}8$	VL	L	VL	H	L	VL	VL	-

Assessment data of each expert in linguistic scale were then converted to triangular fuzzy numbers by using conversion rules given in Table 1. As an example, assessment data of an expert given in Table 4 and Table 5 are redefined as triangular fuzzy numbers in Table 1 and Table 2 in the Appendix. C1 through C15 are critical input factors, C1 through C8 are critical output factors and *l*, *m*, *u* are the lower, medium and upper limit of triangular fuzzy numbers respectively.

Totally 16 assessment matrices (8 for input factors and 8 for output factors) were obtained from experts. Each matrix (also called as initial direct relation matrix) was then normalized by using Eq. (2) and Eq. (3). Then normalized direct relation matrices were obtained by getting average of 8 normalized direct relation matrices by using Eq. (4) for input and output factors separately. Normalized direct relation matrices are shown in Table 3 and Table 4 in the Appendix.

Next, Total relation matrices (Table 5 and Table 6 in the Appendix) were obtained by using Eq. (6). $\tilde{D}_i$ and $\tilde{R}_i$ were calculated. $\tilde{D}_i$ is the sum of the row and $\tilde{R}_i$ is the sum of the column of $\tilde{T}$. Then $\tilde{D}_i$ and $\tilde{R}_i$ were defuzzified separately by using Eq. (7). Then, D_i+R_i and D_i-R_i were calculated. The values of $\tilde{D}_i$, $\tilde{R}_i$, D_i, R_i, D_i+R_i and D_i-R_i are shown in Table 6 for inputs and Table 7 for outputs.

Table 6. Values of $\tilde{D}_i$, $\tilde{R}_i$, D_i, R_i, D_i+R_i and D_i-R_i for Input Factors

Factor	$\tilde{D}_i$			$\tilde{R}_i$			*D*	*R*	*D+R*	*D-R*
	l	***m***	***u***	***l***	***m***	***u***				
1	0.5643	1.1711	2.8896	0.3913	0.8478	2.3079	1.1655	0.9215	2.0870	0.2440
2	0.1611	0.5145	1.7764	0.4155	0.9106	2.3576	0.7099	0.9509	1.6608	-0.2410
3	0.0836	0.2319	1.2905	0.2605	0.5703	1.8835	0.4796	0.7311	1.2107	-0.2515
4	0.8292	1.5889	3.2657	0.3222	0.7974	2.2681	1.3418	0.9144	2.2562	0.4274
5	0.5350	1.1361	2.8835	0.3964	0.9349	2.5081	1.1615	1.0156	2.1771	0.1460
6	0.4811	1.1071	2.8242	0.4471	1.0040	2.6312	1.1500	1.0627	2.2127	0.0873

7	0.6040	1.2490	3.0614	0.5175	1.1457	2.6987	1.2355	1.1090	2.3445	0.1265
8	0.4994	1.0265	2.5361	0.4257	0.8734	2.2721	1.0211	0.9066	1.9276	0.1145
9	0.7635	1.4997	3.3117	0.6128	1.2739	2.9402	1.3493	1.2005	2.5498	0.1488
10	0.6160	1.2933	2.9335	0.4449	0.9319	2.3844	1.2036	0.9572	2.1608	0.2464
11	0.3249	0.8037	2.2827	0.3903	0.8898	2.3819	0.9205	0.9605	1.8810	-0.0399
12	0.1059	0.3858	1.5574	0.3640	0.8571	2.3285	0.6124	0.9405	1.5530	-0.3281
13	0.2284	0.6293	1.9734	0.4903	1.0858	2.7286	0.7915	1.1080	1.8995	-0.3166
14	0.2045	0.5812	1.9073	0.3092	0.6705	2.0628	0.7613	0.8080	1.5694	-0.0467
15	0.0000	0.1141	1.0805	0.2134	0.5391	1.8201	0.3982	0.7153	1.1134	-0.3171

Table 7. Values of $\tilde{D}_i$, $\tilde{R}_i$, D_i, R_i, D_i+R_i and D_i-R_i for Output Factors

Factor	$\tilde{D}_i$			$\tilde{R}_i$			***D***	***R***	***D+R***	***D-R***
	l	***m***	***u***	***l***	***m***	***u***				
1	0.3695	0.7674	2.4300	0.5585	1.2677	3.1778	0.9426	1.2957	2.2383	-0.3530
2	0.2774	0.5798	2.0919	0.6163	1.3487	3.3219	0.7981	1.3514	2.1496	-0.5533
3	0.1018	0.2575	1.5375	0.6761	1.4326	3.4709	0.5644	1.4091	1.9735	-0.8447
4	0.8197	1.5895	3.5137	0.6570	1.3080	3.2704	1.4278	1.3071	2.7350	0.1207
5	0.6321	1.3712	3.2315	0.0000	0.1656	1.3496	1.3236	0.5050	1.8286	0.8185
6	0.6970	1.3561	3.3727	0.2428	0.5225	1.9269	1.3439	0.7355	2.0794	0.6084
7	0.5214	1.0419	2.9354	0.3863	0.7567	2.2522	1.1520	0.8742	2.0262	0.2778
8	0.2100	0.8358	2.3559	0.4918	0.9975	2.6990	0.9939	1.0682	2.0621	-0.0743

Next, the causal diagrams (Figure 1 and Figure 2) were constructed by mapping a dataset of (D+R, D-R).

Figure 1. The Causal Diagram of Input Factors

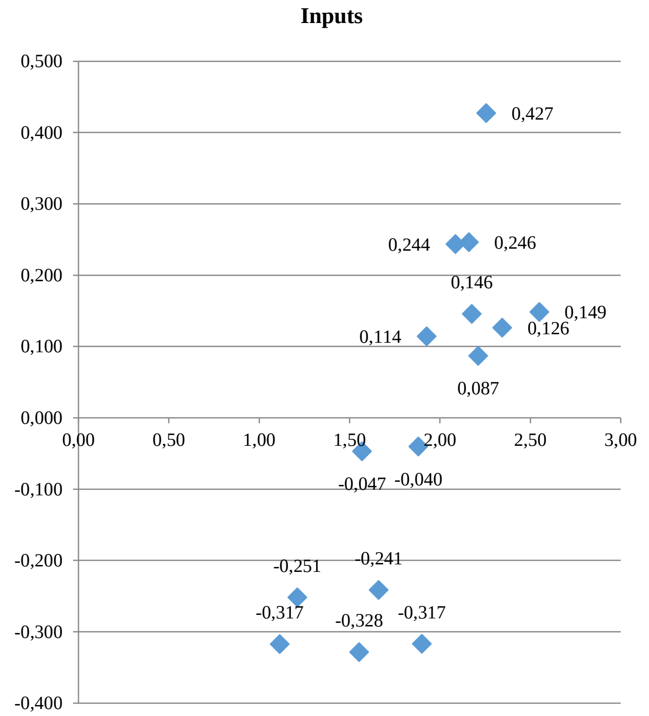
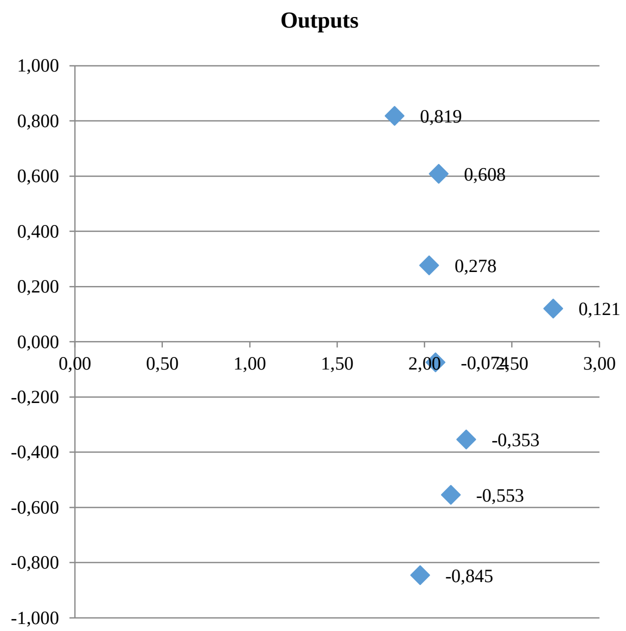


Figure 2. The Causal Diagram of Output Factors



When we look at the causal diagram of input factors (Figure 1), it is easily seen that 8 critical input factors out of 15 belong to cause group. These factors are $C_{in}1$ (Number of Academic Staff), $C_{in}4$ (Number of Undergraduate Students), $C_{in}5$ (Number of MSc.

Students), $C_{in}6$ (Number of PhD. Students), $C_{in}7$ (Total Department Expenditures), $C_{in}8$ (Research Budget), $C_{in}9$ (Department Budget) and $C_{in}10$ (Number of Laboratories). All other factors belong to effect group.

From Figure 2, it is observed that 4 critical output factors out of 8 belong to cause group. These factors are $C_{out}4$ (Number of Papers Published in Academic Journals (SCI, SSCI)), $C_{out}5$ (Number of Undergraduate Awards), $C_{out}6$ (Number of Postgraduate Awards) and $C_{out}7$ (Number of Doctorates Awards). All other factors belong to effect group.

5. Discussions

In this empirical study, the case university departments' efficiency measurements factors are examined through 15 critical input and 8 output factors. According to evaluation results, several implications about efficiency measurement factors can be derived as follows;

Causal diagrams serve us valuable information about effective critical input and output factors in efficiency measurement of university departments. If it is desired to obtain high efficiency in effect group of critical factors, a great deal of attention must be paid to cause group of critical factors. While cause group factors are influencing factors, effect group factors are influenced factors. If we look at the cause and effect relationships in input set, it is clearly seen that number of academic staff, undergraduate students, MSc. students, PhD. students, laboratories, total departments expenditures, research budget and department budget lead to effective efficiency measurement as critical inputs. Similarly, number of academic researches, undergraduate awards, postgraduate awards and doctorates awards are the most important critical outputs for efficiency measurement.

Contrary to common sense, number of papers published in academic journals (SCI, SSCI, peer-reviewed journal, proceedings) were found to be non-critical in efficiency measurement. Analysis of the results shows that number of academic publications varies according to the values of other output factors. Therefore this factor is considered as unnecessary output in efficiency measurement of university departments.

6. Conclusions

Successful efficiency measurement of university departments is very important issue in today's globalized world and depends on paying high attention on critical input and output factors affecting efficiency measurement. In our study, we analyzed 15 critical input and 8 critical output factors by using fuzzy DEMATEL methodology which serves a highly effective structural decision making system for modeling cause and effect relationships. Among 15 critical input factors, Number of Academic Staff, Number of Undergraduate Students, Number of MSc. Students, Number of PhD. Students, Total Department Expenditures, Research Budget, Department Budget and Number of Laboratories and among 8 critical output factors Number of Papers Published in Academic Journals (SCI, SSCI), Number of Undergraduate Awards, Number of Postgraduate Awards and Number of Doctorates Awards are found to be the most important factors that influenced other factors.

This study is the first reference in the literature which uses a fuzzy DEMATEL technique in determination of effective critical input and output factors in successful efficiency measurement of university departments. As an extension of our work, the usage of critical inputs and outputs (determined as cause groups in this study) in the university departments' efficiency measurements via DEA gives researchers more reliable efficiency scores comparing to selection of inputs and outputs conceptually or intuitively.

References

Ahn, T., Arnold, V., Charnes, A., & Cooper, W.W. (1989). DEA and ratio efficiency analyses for public institutions of higher learning in Texas. Research in Governmental and Nonprofit Accounting, 5, 165−185.

Aksakal, E., & Dağdeviren, M. (2010). ANP ve Dematel Yöntemleri ile Personel Seçimi Problemine Bütünleşik Bir Yaklaşım. Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 25(4), 905–913.

Altuntas, S., Selim, H., & Dereli, T. (2014). A fuzzy DEMATEL-based solution approach for facility layout problem: a case study. The International Journal of Advanced Manufacturing Technology, 73(5-8), 749–771.

Arcelus, F. J., & Coleman, D. F. (1997). An efficiency review of university departments. International Journal of Systems Science, 28(7), 721−729.

Athanassopoulos, A. D., & Shale, E. (1997). Assessing the comparative efficiency of higher education institutions in the UK by means of data envelopment analysis. Education Economics, 5(2), 117−134.

Avkiran, N. K. (2001). Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. Socio-Economic Planning Sciences, 35(1), 57−80.

Baykasoglu, A., & Durmusoglu, Z. D. U. (2014). A Hybrid MCDM for Private Primary School Assessment Using DEMATEL Based on ANP and Fuzzy Cognitive Map. International Journal of Computational Intelligence Systems, 7(4), 615–635.

Beasley, J. E. (1990). Comparing university departments. Omega International Journal of Management Science, 18(2), 171–183.

Beasley, J. E. (1995). Determining teaching and research efficiencies. Journal of the Operational Research Society, 46(4), 441–452.

Bellman, R.E., & Zadeh, L.A. (1970). Decision-making in a fuzzy environment, Management Science, 17, 141–164.

Bessent, A. M., Bessent, E. W., Charnes, A., Cooper, W. W., & Thorogood, N. C. (1983). Evaluation of educational program proposals by means of DEA. Educational Administration Quarterly, 19(2), 82–107.

Breu, T. M., & Raab, R. L. (1994). Efficiency and perceived quality of the nation's "Top 25" national universities and national liberal arts colleges: An application of data envelopment analysis to higher education. Socio-Economic Planning Sciences, 28(1), 33−45.

Büyüközkan, G., & Çifçi, G. (2012). A novel hybrid MCDM approach based on fuzzy DEMATEL, fuzzy ANP and fuzzy TOPSIS to evaluate green suppliers. Expert Systems with Applications, 39(3), 3000–3011.

Casu, B., & Thanassoulis, E. (2006). Evaluating cost efficiency in central administrative services in UK universities. Omega, 34, 417–426.

Cave, M., Hanney, S., & Kogan, M. (1991). The Use of performance indicators in higher education: A critical analysis of developing practice. London: Jessica Kingsley Publishers.

Chalmers, D. (2008). Teaching and Learning Quality Indicators in Australian Universities. IMHE/OECD conference. The quality, relevance and impact of higher education. Paris.

Chang, B., Chang, C., & Wu, C. (2011). Fuzzy DEMATEL method for developing supplier selection criteria, Expert Systems with Applications, 38, 1850-1858.

Chou, Y.-C., Sun, C.-C., & Yen, H.-Y. (2012). Evaluating the criteria for human resource for science and technology (HRST) based on an integrated fuzzy AHP and fuzzy DEMATEL approach. Applied Soft Computing, 12(1), 64–71.

Coelli, T. (1996). Assessing the Performance of Australian Universities Using Data Envelopment Analysis. Mimeo. Center for Efficiency and Productivity Analysis. NSW: University of New England.

Coelli, T., Rao, D. S. P., & Battese, G. E. (1998). An Introduction to Efficiency and Productivity Analysis. Kluwer Academic Publishers, Boston, 271 pp.

El Mahgary, S., & Lahdelma, R. (1995). Data envelopment analysis: Visualizing the results. European Journal of Operational Research, 85, 700−710.

Erüz E. (2005). Yeni Mali Yönetim Yapısında Performans Esaslı Bütçeleme, 20. Türkiye Maliye Sempozyumu, 23-27 Mayıs, Denizli, s.61

Flegg, A. T., & Allen, D. O. (2007). Does expansion cause congestion? The case of the older British universities, 1994–2004. Education Economics, 15(1), 75−102.

Flegg, A. T., Allen, D. O., Field, K., & Thurlow, T. W. (2004). Measuring the efficiency and productivity of British universities: An application of DEA and the Malmquist approach. Education Economics, 12(3), 231–249.

Fox, K. J., & Milbourne, R. (1999). What determines research output of academic economists? Economic Record, 75, 256–267.

Friedman, L., & Sinuany-Stern, Z. (1997). Scaling units via the canonical correlation analysis in the data envelopment analysis context. European Journal of Operational Research, 100, 629−637.

Gabus, A., & Fontela, E. (1972). World problems an invitation to further thought within the framework of DEMATEL, Battelle Geneva Research Centre, Switzerland, Geneva.

Gabus, A., & Fontela, E. (1973). Perceptions of the world problematique: Communication procedure, communicating with those bearing collective responsibility (DEMATEL report no. 1). Battelle Geneva Research Centre, Switzerland, Geneva.

Glass, J. C, McKillop, D. G., & Hyndman, N. (1995). Efficiency in the provision of university teaching and research: An empirical analysis of UK universities. Journal of Applied Econometrics, 10(1), 61–72.

Haksever, C., & Muragishi, Y. (1998). Measuring value in MBA programs. Education Economics, 6(1), 11−25.

Hsieh, T. Y., Lu, S. T., & Tzeng, G. H. (2004). Fuzzy MCDM approach for planning and design tenders selection in public office buildings. International Journal of Project Management, 22(7), 573–584.

Johnes J., & Yu L., (2008). Measuring the Research Performance of Chinese Higher Education İnstitutions Using Data Envelopment Analysis. China Economic Review, 19, 679–696

Johnes, J. (1996). Performance assessment in higher education in Britain. European Journal of Operational Research, 89(1), 18–33.

Johnes, J. (2006). DEA and its application to the measurement of efficiency in higher education. Economics of Education Review, 25(3), 273−288.

Johnes, J. (2006). Measuring teaching efficiency in higher education: An application of data envelopment analysis to Economics graduates from UK universities. European Journal of Operational Research, 174, 443−456

Katharaki M., & Katharakis G. (2010). A Comparative Assessment of Greek Universities' Efficiency Using Quantitative Analysis. International Journal of Educational Research, 49(4-5), 115–128

Koksal G., & Nalcacı B. (2006). The Relative Efficiency of Departments at a Turkish Engineering College: A Data Envelopment Analysis. Higher Education, 51, 173–189.

Mavi, R. K., & Shahabi, H. (2015). Using fuzzy DEMATEL for evaluating supplier selection criteria in manufacturing industries. International Journal of Logistics Systems and Management, 22(1), 15–42.

McMillan, M. L., & Datta, D. (1998). The relative efficiency of Canadian universities. Canadian Public Policy, 24(4), 485–511.

Mokhtarian, M. N. (2007). Developing global manager's competencies using the fuzzy DEMATEL method. Expert Systems with Applications, 32(7), 9050-9051

OECD. (2007). Education at a Glance. Retrieved December 12, 2015 from: https://www.oecd.org/education/skills-beyond-school/40701218.pdf.

Organ, A. (2013). Bulanık Dematel Yöntemiyle Makine Seçimini Etkileyen Kriterlerin Değerlendirilmesi. Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 22(1), 157–172.

Ortíz, M. A., Felizzola, H. A., & Isaza, S. (2015). A contrast between DEMATEL-ANP and ANP methods for six sigma project selection: a case study in healthcare industry. BMC Medical Informatics and Decision Making, 15(Suppl 3), S3.

Oruc K.O. (2008). Efficiency Measurements in Fuzzy Environment by Using Data Envelopment Analysis and an Application at Universities, Süleyman Demirel University, Department of Business Administration, Ph.D.

Pereira, A. P., & Tavares, A. F. (2002). O Uso de Indicadores de Performance na Avaliac¸a˜o de Ensino Superior. Revista de Administrac¸a˜o e Polı´ticas Pu´blicas, 3(1–2), 82–94.

Preeti T., Shiv P.Y., & Singh S.P. (2009).Relative Performance of Academic Departments Using DEA With Sensitivity Analysis, Evaluation and Program Planning, 32(2), 168-177

Salerno, C. (2006). Using data envelopment analysis to improve estimates of higher education institution's per student education costs. Education Economics, 14(3), 281–295.

Sener, Z. (2016). Evaluating Ship Selection Criteria for Maritime Transportation. Journal of Advanced Management Science, 4(4), 325–328.

Stevens, P. A. (2001). The Determinants of economic efficiency in English and Welsh universities. London: National Institute of Economic and Social Research. Discussion Paper no. 185.

Tsai, S.-B., Chien, M.-F., Xue, Y., Li, L., Jiang, X., Chen, Q., Zhou, J., & Wang, L. (2015). Using the Fuzzy DEMATEL to Determine Environmental Performance: A Case of Printed Circuit Board Industry in Taiwan. PLOS ONE, 10(6), e0129153.

Tseng, M.-L. (2009). A causal and effect decision making model of service quality expectation using grey-fuzzy DEMATEL approach. Expert Systems with Applications, 36(4), 7738–7748.

Vinodh, S., & Swarnakar, V. (2015). Lean Six Sigma project selection using hybrid approach based on fuzzy DEMATEL–ANP–TOPSIS. International Journal of Lean Six Sigma, 6(4), 313–338.

Ward, D. (2007). Academic values, institutional management, and public policies. Higher Education Management and Policy, 19(2), 1–12.

Worthington, A. (2001). An Empirical survey of frontier efficiency measurement techniques in education. Education Economics, 9(3), 245–268.

Worthington, A. C., & Lee, B. L. (2008). Efficiency, technology and productivity change in Australian universities 1998–2003. Economics of Education Review, 27(3), 285−298.

Wu, W.-W., & Lee, Y.-T. (2007). Developing global managers' competencies using the fuzzy DEMATEL method. Expert Systems with Applications, 32, 499–507.

YÖK (Yüksek Öğretim Kurumu). (2007). Türkiyenin Yüksek Öğrenim Stratejisi. (Yayın No: 2007-1), Ankara, YÖK.

Zadeh, L.A. (1965). Fuzzy sets, Information and Control, 8, 338–353.

Zhou, Q., Huang, W., & Zhang, Y. (2011). Identifying critical success factors in emergency management using a fuzzy DEMATEL method, Safety Science, 49, 243-252.

Appendix.

Table 1. Assessment Data of an Expert for Input Factors in Triangular Fuzzy Numbers

	C_{in} 1			C_{in} 2			C_{in} 3			C_{in} 4			C_{in} 5			C_{in} 6			C_{in} 7			C_{in} 8			C_{in} 9			C_{in} 10			C_{in} 11			C_{in} 12			C_{in} 13			C_{in} 14			C_{in} 15		
	l	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*
C_{in} 1	0	0	0	0.5	0.75	1	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0	0	0.25	0	0.25	0.5	0	0	0.25	0	0.25	0.5	0.75	1	1	0	0.25	0.5
C_{in} 2	0	0	0.25	0	0	0	0.25	0.5	0.75	0	0.25	0.5	0.25	0.5	0.75	0.25	0.5	0.75	0	0.25	0.5	0	0.25	0.5	0	0.25	0.5	0	0	0.25	0	0	0.25	0	0	0.25	0	0.25	0.5	0	0	0.25	0	0	0.25
C_{in} 3	0	0	0.25	0	0	0.25	0	0	0	0	0	0.25	0	0	0.25	0	0	0.25	0.25	0.5	0.75	0	0	0.25	0.25	0.5	0.75	0	0	0.25	0	0	0.25	0	0	0.25	0.25	0.5	0.75	0	0	0.25	0	0	0.25
C_{in} 4	0.75	1	1	0.75	1	1	0.5	0.75	1	0	0	0	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.25	0.5	0.75	0.75	1	1	0.75	1	1	0.75	1	1	0.75	1	1	0.75	1	1	0.5	0.75	1	0.25	0.5	0.75
C_{in} 5	0.75	1	1	0.5	0.75	1	0.5	0.75	1	0	0	0.25	0	0	0	0.75	1	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.75	1	1	0	0	0.25
C_{in} 6	0.5	0.75	1	0.25	0.5	0.75	0.25	0.5	0.75	0	0.25	0.5	0	0.25	0.5	0	0	0	0.25	0.5	0.75	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0	0	0.25
C_{in} 7	0.5	0.75	1	0.5	0.75	1	0	0	0.25	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0	0	0	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0	0	0.25	0	0.25	0.5
C_{in} 8	0.5	0.75	1	0.5	0.75	1	0	0	0.25	0	0	0.25	0.5	0.75	1	0.5	0.75	1	0.75	1	1	0	0	0	0.75	1	1	0.5	0.75	1	0.5	0.75	1	0	0.25	0.5	0	0	0.25	0	0	0.25	0	0	0.25
C_{in} 9	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.75	1	1	0.75	1	1	0	0	0	0.75	1	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.25	0.5	0.75	0.75	1	1
C_{in} 10	0.25	0.5	0.75	0.25	0.5	0.75	0	0	0.25	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.75	1	1	0.75	1	1	0.75	1	1	0	0	0	0	0.25	0.5	0	0.25	0.5	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1
C_{in} 11	0	0	0.25	0	0	0.25	0	0	0.25	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0	0	0.25	0.75	1	1	0.25	0.5	0.75	0	0	0	0.25	0.5	0.75	0.5	0.75	1	0.5	0.75	1	0.25	0.5	0.75
C_{in} 12	0	0	0.25	0	0	0.25	0	0	0.25	0	0.25	0.5	0	0.25	0.5	0	0.25	0.5	0.25	0.5	0.75	0	0	0.25	0.25	0.5	0.75	0	0	0.25	0	0	0.25	0	0	0	0.5	0.75	1	0	0	0.25	0	0	0.25
C_{in} 13	0	0	0.25	0	0	0.25	0	0	0.25	0.5	0.75	1	0.5	0.75	1	0.5	0.75	1	0	0.25	0.5	0	0	0.25	0.25	0.5	0.75	0	0.25	0.5	0	0.25	0.5	0.25	0.5	0.75	0	0	0	0	0	0.25	0	0	0.25
C_{in} 14	0.5	0.75	1	0.25	0.5	0.75	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0.25	0.5	0.75	0	0	0.25	0.25	0.5	0.75	0.25	0.5	0.75	0.25	0.5	0.75	0.5	0.75	1	0	0.25	0.5	0	0	0	0.25	0.5	0.75
C_{in} 15	0	0.25	0.5	0	0.25	0.5	0	0	0.25	0	0.25	0.5	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0

Table 2. Assessment Data of an Expert for Output Factors in Triangular Fuzzy Numbers

	C_{out} 1			C_{out} 2			C_{out} 3			C_{out} 4			C_{out} 5			C_{out} 6			C_{out} 7			C_{out} 8		
	l	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*
C_{out} 1	0	0	0	0.75	1	1	0.5	0.75	1	0.25	0.5	0.75	0	0	0.25	0	0	0.25	0	0	0.25	0	0.25	0.5
C_{out} 2	0.5	0.75	1	0	0	0	0.25	0.5	0.75	0.5	0.75	1	0	0	0.25	0	0.25	0.5	0	0	0.25	0	0	0.25
C_{out} 3	0.25	0.5	0.75	0	0.25	0.5	0	0	0	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25	0	0	0.25
C_{out} 4	0.5	0.75	1	0.5	0.75	1	0.75	1	1	0	0	0	0	0.25	0.5	0.5	0.75	1	0.5	0.75	1	0.75	1	1
C_{out} 5	0.25	0.5	0.75	0	0.25	0.5	0.25	0.5	0.75	0.25	0.5	0.75	0	0	0	0.75	1	1	0.5	0.75	1	0	0.25	0.5
C_{out} 6	0.5	0.75	1	0.5	0.75	1	0.75	1	1	0.5	0.75	1	0	0	0.25	0	0	0	0.75	1	1	0.5	0.75	1
C_{out} 7	0.5	0.75	1	0.75	1	1	0.5	0.75	1	0.25	0.5	0.75	0	0	0.25	0	0	0.25	0	0	0	0.5	0.75	1
C_{out} 8	0	0.25	0.5	0.25	0.5	0.75	0	0.25	0.5	0.5	0.75	1	0.25	0.5	0.75	0	0.25	0.5	0	0.25	0.5	0	0	0

Table 3. The Normalized Direct-Relation Matrix for Output Factors

	C_{out} 1			C_{out} 2			C_{out} 3			C_{out} 4			C_{out} 5			C_{out} 6			C_{out} 7			C_{out} 8		
	l	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*
C_{out} 1	0.0000	0.0000	0.0000	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0385	0.0769	0.1154
C_{out} 2	0.0385	0.0769	0.1154	0.0000	0.0000	0.0000	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385
C_{out} 3	0.0385	0.0769	0.1154	0.0385	0.0769	0.1154	0.0000	0.0000	0.0000	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385
C_{out} 4	0.1154	0.1538	0.1538	0.1154	0.1538	0.1538	0.1154	0.1538	0.1538	0.0000	0.0000	0.0000	0.0000	0.0385	0.0769	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.1154	0.1538	0.1538
C_{out} 5	0.0385	0.0769	0.1154	0.0385	0.0769	0.1154	0.0385	0.0769	0.1154	0.0385	0.0769	0.1154	0.0000	0.0000	0.0000	0.1154	0.1538	0.1538	0.1154	0.1538	0.1538	0.0385	0.0769	0.1154
C_{out} 6	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0000	0.0000	0.0385	0.0000	0.0000	0.0000	0.1154	0.1538	0.1538	0.0769	0.1154	0.1538
C_{out} 7	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0769	0.1154	0.1538	0.0000	0.0000	0.0385	0.0000	0.0000	0.0385	0.0000	0.0000	0.0000	0.0769	0.1154	0.1538
C_{out} 8	0.0000	0.0385	0.0769	0.0000	0.0385	0.0769	0.0000	0.0385	0.0769	0.1154	0.1538	0.1538	0.0000	0.0385	0.0769	0.0000	0.0385	0.0769	0.0000	0.0385	0.0769	0.0000	0.0000	0.0000

Table 4. The Normalized Direct-Relation Matrix for Input Factors

	C_{in} 1			C_{in} 2			C_{in} 3			C_{in} 4			C_{in} 5			C_{in} 6			C_{in} 7			C_{in} 8			C_{in} 9			C_{in} 10			C_{in} 11			C_{in} 12			C_{in} 13			C_{in} 14			C_{in} 15		
	l	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*
C_{in} 1	0.0000	0.0000	0.0000	0.0545	0.0727	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545
C_{in} 2	0.0000	0.0000	0.0182	0.0000	0.0000	0.0000	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545	0.0000	0.0182	0.0364	0.0000	0.0182	0.0364	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0182	0.0364	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182
C_{in} 3	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0000	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182
C_{in} 4	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0000	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545
C_{in} 5	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0000	0.0000	0.0000	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0182
C_{in} 6	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0000	0.0182	0.0364	0.0000	0.0182	0.0364	0.0000	0.0000	0.0000	0.0182	0.0364	0.0545	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0182
C_{in} 7	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0182	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0000	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545
C_{in} 8	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0545	0.0727	0.0727	0.0000	0.0000	0.0000	0.0545	0.0727	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0182	0.0364	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182
C_{in} 9	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0000	0.0000	0.0000	0.0545	0.0727	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545	0.0364	0.0545	0.0727
C_{in} 10	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0545	0.0727	0.0727	0.0000	0.0000	0.0000	0.0000	0.0182	0.0364	0.0000	0.0182	0.0364	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727
C_{in} 11	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0000	0.0000	0.0000	0.0182	0.0364	0.0364	0.0545	0.0727	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545
C_{in} 12	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0182	0.0364	0.0000	0.0182	0.0364	0.0000	0.0182	0.0364	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0000	0.0364	0.0545	0.0727	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182
C_{in} 13	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0364	0.0545	0.0727	0.0000	0.0182	0.0364	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0000	0.0182	0.0364	0.0000	0.0182	0.0364	0.0182	0.0364	0.0545	0.0000	0.0000	0.0000	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182
C_{in} 14	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0000	0.0000	0.0182	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0182	0.0364	0.0545	0.0364	0.0545	0.0727	0.0000	0.0182	0.0364	0.0000	0.0000	0.0000	0.0000	0.0182	0.0364
C_{in} 15	0.0000	0.0182	0.0364	0.0000	0.0182	0.0364	0.0000	0.0000	0.0182	0.0000	0.0182	0.0364	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0182	0.0000	0.0000	0.0000

Table 5. Total Relation Matrix for Input Factors

	C_{in} 1			C_{in} 2			C_{in} 3			C_{in} 4			C_{in} 5			C_{in} 6			C_{in} 7			C_{in} 8			C_{in} 9			C_{in} 10			C_{in} 11			C_{in} 12			C_{in} 13			C_{in} 14			C_{in} 15		
	l	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*
C_{in} 1	0.0123	0.0381	0.1365	0.0658	0.1102	0.2075	0.0444	0.0798	0.1805	0.0439	0.0849	0.1989	0.0461	0.0914	0.2139	0.0477	0.0946	0.2214	0.0499	0.1009	0.2256	0.0477	0.0902	0.2017	0.0525	0.1064	0.2398	0.0141	0.0430	0.1600	0.0305	0.0755	0.1931	0.0117	0.0387	0.1558	0.0308	0.0803	0.2097	0.0438	0.0804	0.1883	0.0231	0.0569	0.1569
C_{in} 2	0.0039	0.0183	0.1025	0.0040	0.0183	0.0856	0.0384	0.0652	0.1379	0.0205	0.0496	0.1325	0.0037	0.0365	0.1259	0.0039	0.0378	0.1303	0.0228	0.0580	0.1494	0.0214	0.0529	0.1344	0.0230	0.0601	0.1577	0.0044	0.0203	0.1053	0.0039	0.0192	0.1052	0.0032	0.0178	0.1026	0.0042	0.0390	0.1332	0.0018	0.0117	0.0909	0.0019	0.0099	0.0828
C_{in} 3	0.0019	0.0078	0.0786	0.0019	0.0080	0.0797	0.0011	0.0046	0.0488	0.0024	0.0094	0.0794	0.0026	0.0101	0.0852	0.0026	0.0105	0.0882	0.0199	0.0448	0.1217	0.0023	0.0086	0.0778	0.0201	0.0454	0.1281	0.0023	0.0095	0.0811	0.0019	0.0086	0.0810	0.0022	0.0091	0.0803	0.0202	0.0453	0.1238	0.0008	0.0045	0.0705	0.0013	0.0056	0.0662
C_{in} 4	0.0676	0.1161	0.2196	0.0693	0.1206	0.2229	0.0466	0.0858	0.1925	0.0151	0.0486	0.1500	0.0520	0.1071	0.2332	0.0539	0.1110	0.2414	0.0568	0.1189	0.2459	0.0361	0.0865	0.2017	0.0775	0.1426	0.2620	0.0702	0.1220	0.2250	0.0675	0.1193	0.2248	0.0672	0.1182	0.2219	0.0736	0.1332	0.2488	0.0484	0.0916	0.2044	0.0273	0.0671	0.1716
C_{in} 5	0.0455	0.0864	0.2021	0.0287	0.0721	0.1885	0.0242	0.0576	0.1602	0.0096	0.0350	0.1516	0.0111	0.0401	0.1471	0.0479	0.0951	0.2224	0.0327	0.0846	0.2099	0.0475	0.0895	0.2009	0.0521	0.1062	0.2405	0.0472	0.0913	0.2077	0.0455	0.0893	0.2075	0.0446	0.0877	0.2042	0.0484	0.0975	0.2278	0.0440	0.0806	0.1888	0.0060	0.0230	0.1243
C_{in} 6	0.0439	0.0854	0.1990	0.0277	0.0717	0.1859	0.0233	0.0571	0.1580	0.0092	0.0519	0.1660	0.0107	0.0578	0.1799	0.0111	0.0420	0.1512	0.0315	0.0838	0.2070	0.0458	0.0879	0.1976	0.0503	0.1050	0.2368	0.0456	0.0902	0.2044	0.0439	0.0882	0.2042	0.0430	0.0867	0.2011	0.0467	0.0965	0.2243	0.0425	0.0794	0.1859	0.0058	0.0233	0.1230
C_{in} 7	0.0480	0.0922	0.2118	0.0488	0.0952	0.2149	0.0090	0.0286	0.1366	0.0462	0.0910	0.2106	0.0488	0.0984	0.2264	0.0505	0.1019	0.2343	0.0166	0.0538	0.1676	0.0498	0.0954	0.2114	0.0550	0.1120	0.2509	0.0498	0.0966	0.2170	0.0480	0.0945	0.2168	0.0464	0.0919	0.2124	0.0514	0.1033	0.2380	0.0107	0.0337	0.1482	0.0249	0.0604	0.1648
C_{in} 8	0.0471	0.0871	0.1889	0.0479	0.0900	0.1917	0.0081	0.0248	0.1172	0.0107	0.0345	0.1375	0.0472	0.0911	0.1995	0.0489	0.0944	0.2065	0.0679	0.1155	0.2083	0.0152	0.0410	0.1223	0.0696	0.1196	0.2212	0.0491	0.0911	0.1932	0.0470	0.0889	0.1930	0.0097	0.0509	0.1548	0.0139	0.0443	0.1609	0.0095	0.0293	0.1275	0.0075	0.0239	0.1136
C_{in} 9	0.0507	0.0992	0.2219	0.0516	0.1025	0.2252	0.0446	0.0824	0.1940	0.0481	0.0963	0.2197	0.0515	0.1047	0.2359	0.0533	0.1084	0.2442	0.0731	0.1323	0.2481	0.0698	0.1181	0.2200	0.0245	0.0711	0.1966	0.0699	0.1203	0.2271	0.0503	0.1011	0.2269	0.0484	0.0984	0.2230	0.0542	0.1115	0.2504	0.0295	0.0715	0.1894	0.0438	0.0819	0.1894
C_{in} 10	0.0331	0.0803	0.1919	0.0331	0.0825	0.1942	0.0086	0.0283	0.1306	0.0460	0.0908	0.2029	0.0496	0.0991	0.2181	0.0514	0.1026	0.2257	0.0699	0.1243	0.2277	0.0683	0.1142	0.2049	0.0724	0.1299	0.2421	0.0165	0.0485	0.1437	0.0153	0.0646	0.1792	0.0139	0.0621	0.1758	0.0501	0.1026	0.2287	0.0451	0.0846	0.1915	0.0426	0.0789	0.1765
C_{in} 11	0.0074	0.0278	0.1267	0.0069	0.0274	0.1273	0.0040	0.0160	0.1039	0.0245	0.0608	0.1583	0.0256	0.0642	0.1687	0.0266	0.0665	0.1746	0.0265	0.0685	0.1753	0.0086	0.0281	0.1240	0.0455	0.0895	0.2033	0.0440	0.0830	0.1808	0.0067	0.0279	0.1111	0.0075	0.0471	0.1460	0.0445	0.0865	0.1960	0.0238	0.0571	0.1489	0.0226	0.0535	0.1380
C_{in} 12	0.0020	0.0135	0.0917	0.0020	0.0132	0.0924	0.0011	0.0085	0.0775	0.0031	0.0313	0.1084	0.0033	0.0334	0.1161	0.0034	0.0346	0.1202	0.0200	0.0508	0.1361	0.0024	0.0138	0.0905	0.0206	0.0531	0.1444	0.0024	0.0158	0.0948	0.0020	0.0148	0.0947	0.0026	0.0155	0.0763	0.0386	0.0699	0.1561	0.0009	0.0095	0.0827	0.0013	0.0079	0.0754
C_{in} 13	0.0067	0.0234	0.1130	0.0055	0.0224	0.1129	0.0043	0.0156	0.0952	0.0385	0.0710	0.1589	0.0400	0.0755	0.1697	0.0415	0.0782	0.1757	0.0061	0.0450	0.1433	0.0060	0.0235	0.1113	0.0255	0.0662	0.1703	0.0072	0.0438	0.1336	0.0067	0.0420	0.1329	0.0247	0.0601	0.1490	0.0078	0.0298	0.1126	0.0055	0.0199	0.1042	0.0022	0.0130	0.0908
C_{in} 14	0.0211	0.0509	0.1386	0.0221	0.0538	0.1419	0.0028	0.0117	0.0900	0.0043	0.0206	0.1086	0.0042	0.0212	0.1157	0.0044	0.0220	0.1197	0.0236	0.0596	0.1556	0.0049	0.0194	0.1056	0.0239	0.0613	0.1644	0.0219	0.0531	0.1418	0.0211	0.0519	0.1417	0.0388	0.0695	0.1572	0.0057	0.0416	0.1411	0.0028	0.0133	0.0784	0.0029	0.0313	0.1070
C_{in} 15	0.0000	0.0213	0.0851	0.0000	0.0227	0.0870	0.0000	0.0042	0.0606	0.0000	0.0215	0.0848	0.0000	0.0043	0.0729	0.0000	0.0044	0.0754	0.0000	0.0051	0.0771	0.0000	0.0042	0.0680	0.0000	0.0056	0.0821	0.0000	0.0034	0.0692	0.0000	0.0039	0.0697	0.0000	0.0032	0.0680	0.0000	0.0046	0.0771	0.0000	0.0033	0.0633	0.0000	0.0024	0.0398

Table 6. Total Relation Matrix for Output Factors

	C_{out} 1			C_{out} 2			C_{out} 3			C_{out} 4			C_{out} 5			C_{out} 6			C_{out} 7			C_{out} 8		
	l	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*	*l*	*m*	*u*
C_{out} 1	0.0192	0.0647	0.2690	0.0940	0.1741	0.4162	0.0975	0.1803	0.4305	0.0927	0.1678	0.4091	0.0000	0.0109	0.1542	0.0071	0.0254	0.2017	0.0080	0.0294	0.2241	0.0511	0.1149	0.3252
C_{out} 2	0.0536	0.1235	0.3376	0.0186	0.0565	0.2458	0.0934	0.1657	0.3922	0.0850	0.1473	0.3685	0.0000	0.0071	0.1368	0.0065	0.0195	0.1793	0.0073	0.0225	0.1992	0.0129	0.0376	0.2326
C_{out} 3	0.0413	0.0914	0.2737	0.0428	0.0947	0.2832	0.0073	0.0266	0.1895	0.0068	0.0242	0.2158	0.0000	0.0014	0.1086	0.0005	0.0035	0.1366	0.0006	0.0040	0.1518	0.0025	0.0117	0.1783
C_{out} 4	0.1465	0.2694	0.5260	0.1519	0.2791	0.5441	0.1575	0.2890	0.5629	0.0523	0.1409	0.4043	0.0000	0.0530	0.2364	0.0809	0.1489	0.3667	0.0903	0.1718	0.4075	0.1402	0.2373	0.4657
C_{out} 5	0.0731	0.1915	0.4656	0.0758	0.1983	0.4817	0.0786	0.2054	0.4983	0.0777	0.1949	0.4760	0.0000	0.0140	0.1526	0.1214	0.1850	0.3529	0.1354	0.2134	0.3922	0.0700	0.1687	0.4123
C_{out} 6	0.1096	0.2214	0.5105	0.1137	0.2293	0.5281	0.1179	0.2375	0.5463	0.1165	0.2253	0.5218	0.0000	0.0162	0.1980	0.0090	0.0360	0.2216	0.1254	0.1954	0.3944	0.1049	0.1950	0.4520
C_{out} 7	0.0983	0.1919	0.4594	0.1019	0.1987	0.4753	0.1057	0.2058	0.4917	0.1045	0.1953	0.4696	0.0000	0.0140	0.1782	0.0080	0.0312	0.2328	0.0090	0.0360	0.2216	0.0941	0.1690	0.4068
C_{out} 8	0.0169	0.1139	0.3360	0.0175	0.1180	0.3476	0.0182	0.1222	0.3596	0.1214	0.2122	0.4053	0.0000	0.0491	0.1847	0.0093	0.0729	0.2353	0.0104	0.0842	0.2614	0.0162	0.0633	0.2260

Available online at www.alphanumericjournal.com

alphanumeric journal

The Journal of Operations Research, Statistics, Econometrics and Management Information Systems

Volume 4, Issue 1, 2016

2016.04.01.OR.05

ENERGY SAVING IN CONTINUOUS ANNEALING LINE USING HEATING OPTIMIZATION

Ömer Faruk SEZER *

Faculty of Management, Sakarya University, Sakarya, Turkey

Erman COŞKUN †

Prof. Dr., Faculty of Management, Sakarya University, Sakarya, Turkey

Received: 17 February 2016

Accepted: 29 June 2016

Abstract

Energy consumption of Iron and Steel Industry sector in Turkey has the highest share in final energy consumption. In the globalized world, day by day, worsening of the conditions of competition and negative environmental pressures, more efficient energy usage is come to the forefront. Growing industries can survive by providing efficient energy consumption and effective energy follow-up. The failure of businesses is become inevitable with ineffective and inefficient usage of energy resources used in the production. Energy savings in the continuous annealing line which is the subject of this article is an important issue and it is based on many parameters. Measuring strip's temperature is so important to continuous annealing line control in the annealing automation system. Increasing or decreasing the oven temperature provides the process control with the adjustment of the heater which is specific heating capacity in these furnaces. Each strip annealing temperature is different which depends on the strip quality. Also strip thickness, strip width and the line speed are the other factors of the oven temperature conditions. It needs to realize the production of an optimal quality ranking to increase the line speed. This improves the production level and because of the oven temperature changes in the order that causes the reduction of costs due to energy savings, optimized production is achieved. Finding the most suitable production sequence is modeled by fuzzy goal programming functions used the above parameters were investigated.

Keywords: Continuous Annealing Line, Energy Saving, Goal Programming

SÜREKLİ TAVLAMA HATLARI ISITMA OPTİMİZASYONU İLE ENERJİ TASARRUFU SAĞLANMASI

Özet

Demir ve Çelik Sanayi enerji tüketimi ülkemizde nihai enerji tüketimlerinde en yüksek paya sahiptir. Küreselleşen dünyada rekabet koşullarının ve çevre koşullarına uyum şartlarının gün geçtikçe daha da ağırlaşması enerjinin etkin kullanımını ön plana çıkarmaktadır. Gelişen sanayiler enerjiyi etkin takip ederek ve enerji tüketimini verimli kullanarak ayakta kalabileceklerdir. Enerji kaynaklarını üretimde etkin ve verimli kullanmayan işletmelerin başarısızlığa uğraması kaçınılmazdır. Bu tezin konusu olan sürekli tavlama hatlarındaki enerji tasarrufu önemli bir konudur ve birçok parametreye bağlıdır. Tavlama otomasyon sisteminde şerit sıcaklığının ölçülmesi sürekli tavlama hat kontrolünde oldukça önemlidir. Belirli ısıtma kapasitesine sahip ısıtıcıların değerlerinin arttırılıp azaltılması ile fırın içi sıcaklarının dolayısıyla şerit sıcaklıklarının kontrolü sağlanır. Her bir şerit tavlama sıcaklığı şerit kalitesine bağlı olarak farklıdır. Ayrıca fırın sıcaklık değeri şerit boyutlarına ve hat hızına da bağlı olarak değişir. Hat hızını arttırmak için en uygun kalite sırasında üretimi gerçekleştirmek gerekir. Üretim seviyesinin arttırılması ve sıcaklık değişimlerinin en uygun olduğu sırada ürünlerin üretilmesi enerji tasarrufu oluşturmaktadır; bu da maliyetlerin azalmasına neden olur. Yukarıdaki parametrelerin modellendiği bulanık mantık hedef programlama fonksiyonları ile en uygun üretim sıralamasının bulunması araştırılmıştır.

Anahtar Kelimeler: Sürekli Tavlama Hattı, Enerji Tasarrufu, Bulanık Hedef Programlama

* omer.sezer@ogr.sakarya.edu.tr

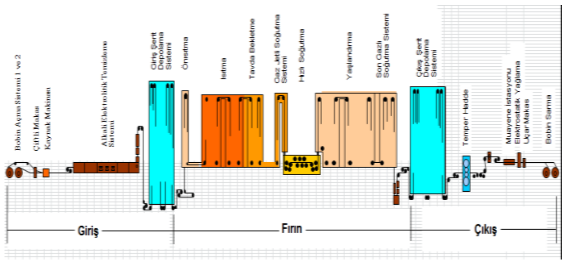
† ermanc@sakarya.edu.tr (Corresponding author)

1. GİRİŞ

Mühendislikte sıcak yahut soğuk haldeki metal malzemeye basit biçimler vermekte kullanılan işleme haddeleme işlemi denir (Sert,2105). Özellikle çeliğin haddelenmesi, sadece biçim vermek değil, aynı zamanda çeliğin mekanik özelliklerinin geliştirilmesini de sağlar. Haddeleme işleminde metal, aynı hızda fakat birbirine ters yönde dönen ve aralarındaki mesafe malzemenin kalınlığından biraz az olan iki merdane hadde arasından geçirilir. Merdaneler ekseriya yüksek kaliteli çelikten yapılmıştır. Haddeden çekme işi, çamaşırları eski çamaşır makinesinin sıkıcısının silindirleri arasından geçirmekten farklı değildir. Çamaşırda düşünülen şey suyu sıkıp çıkarmaktır, madenlerde ise sıkıştırıp daha ince bir hale getirmektir (Sert,2105).

Sürekli tavlama hattı, soğuk haddeleme sürecinin en son işlemi olarak, soğuk haddelenen ürünlerin performansını iyileştirmede önemli bir rol alır. Nippon Steel Kimitsu Fabrikası, Ekim 1972 yılında ilk sürekli tavlama hattını kurduktan sonra, yüksek verimlilik, düşük maliyet, yüksek kalite ve çeşitlendirme avantajları sunan bu işlem tüm dünyada uygulanmaya başlamıştır [Şekil-1].

Şekil 1. Sürekli Tavlama Hattı Akışı



Çelik, yüksek basınçta ince bir şerit haline haddelendiği zaman sertleşir ve en son ürünler olan otomobil parçaları, bilgisayar kasaları, yapı elemanları ve benzerleri gibi ürün haline getirilmeye çalışıldığı zaman başarılı olunamaz. Sürekli tavlama hattı fırın bölümünde; sac bünyesinde soğuk haddeleme sırasında oluşan içyapı bozukluğu ve iç gerilme birikiminin yeniden kristalleşme sıcaklığına kadar tavlanıp giderilmesi amaçlanmaktadır. Saca istenilen fiziksel özellikleri kazandırmak için çeşitli ısıtma ve soğutma işlemleri uygulanır. Bu işlemler, sacın işleme performansını arttırır. Alan çalışması yapacağımız işletmede ısıtma ve soğutma işlemlerinin değişimine neden olan 5 çeşit kalite kullanılmaktadır. Bu kaliteler çeşitli mekaniksel özelliklerine bağlı olarak CQ, DQ1, DQ2, DDQ ve EDDQ adlarını alırlar. Bunların karbon, mangan ve kükürt miktarları sırasıyla yüksek değerden düşük değere sıralanırlar. Ayrıca bu değerlere bağlı olarak malzeme akma ve çekme mukavemetleri değişir. Bu değerlere göre de değişik işlerde kullanılırlar. CQ kalite ticari kalite olup soğuk haddeleme ürünleri için genel bir uygulamadır. Sırasıyla basit uygulamalardan daha karmaşık şekilli ürünlerin üretilmesinde kullanılırlar. EDDQ kalite eşya ve otomotiv sanayinde çok fazla derin çekme ve son derece karmaşık şekilli parçaların imalatında kullanılır. Bunların tavlama sıcaklıkları ve tavlamada kalma süreleri farklı farklıdır. Yine sırasıyla tavlama sıcaklıkları 720 ± 30 dereceden başlayıp en son EDDQ kalite için 830 ± 20 dereceye kadar çıkar ve hıza bağlı olarak tavlama fırınında şeridin belirli sürenin üstünde kalması gerekmektedir. Kalitelere göre sıcaklık ve maruz kalma süre dağılımları Tablo-1'de gösterilmiştir.

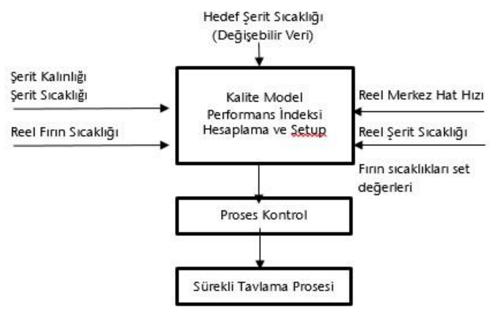
Tablo 1. Kaliteye göre şerit tavlama sıcaklıkları

Ürün Kalitesi	Şerit Sıcaklıkları[°C]	Zaman[saniye]
CQ	720 ± 30	≥ 40
DQ-1	770 ± 30	≥ 40
DQ-2	780 ± 30	≥ 40
DDQ	800 ± 30	≥ 40
EDDQ	850 ± 30	≥ 40

Fırının içinde şeride olan ısı transferi fırın sıcaklığına, şeridin ebatlarına ve hat hızına doğrudan bağlıdır. Bu nedenlerle malzeme kalite performansını etkileyen faktörler arasında, şerit sıcaklığı, kalınlığı, genişliği, fırın sıcaklığı ve hat hızı anahtar durumdadır [Şek.2]. Sürekli tavlama fırınının sıcaklık kontrol sonucunu geleneksel kontrol yöntemini kullanarak elde etmek zordur. Tavlama işlemi kontrolü zor ve karmaşık bir endüstriyel işlem olup, çok iyi bir performans ile iyi bir ısı kontrol yöntemi büyük önem taşımaktadır (Guo, Zhang, You, Chen ve Zhang (2009)).

Her fırın çıkışında fırını ne kadar doğru kontrol ederseniz, ürün kalitesi o kadar etkilenir (Ming, Datai, Jiangang (2008)).

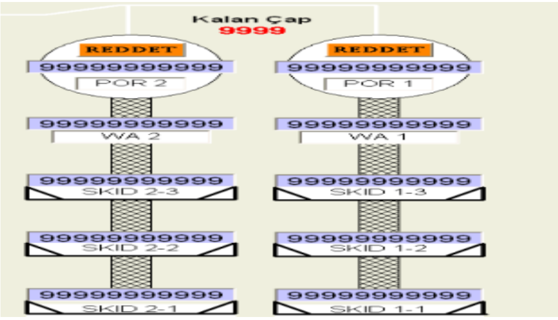
Şekil 2. Fırın Optimizasyon Yapısı



Bunun yanı sıra günümüzde rakiplerle yarışabilmek için yukarıda sözü edilen kalite koşullarının yerine getirilmesinin yanında en ucuz maliyet ile bu işlemin yapılması gerekmektedir.

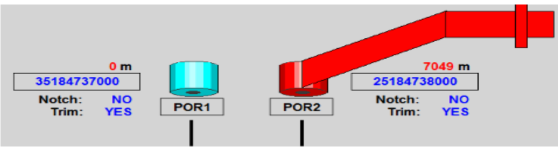
Bu çalışmada hatta girecek olan şerit için en uygun tavlama sıcaklığının bulunması amaçlanmıştır. Şekil-3'de tavlama hatlarında kullanılan hatta girecek olan bobinlerin istiflendiği konveyör görülmektedir.

Şekil 3. Sürekli Tavlama Hattı Giriş Konveyörü



Şekil 4.'de ise bobinin hatta veriliş tarzı gösterilmektedir. Her hatta 2. Mandrel'de bobin var ise 1. Mandrel'deki bobin 2. Mandrel'deki bobinin arkasına kaynak yapılması için bekletilmektedir. Tam tersi de geçerlidir.

Şekil 4. Sürekli Tavlama Hattı Mandrel Görünüşü



Bu çalışmada Şekil 3.'de görülen her bobin için en uygun tavlama sıcaklıklarının bulunması amaçlanmıştır. Bu bir dinamik yapıdır ve gelen her bobin için tekrar hesaplanmaktadır.

Sürekli tavlama ile ilgili çalışmalar literatür taramasında gösterilmiştir. Bu çalışmalarda genellikle matematiksel model üzerinde durulmuş, kalite geçişleri ile ilgili bir çalışma yapılmamıştır. Bu çalışmada rekabet için önemli olan üretim kapasitesini sipariş kalitelerine bağlı olarak maksimize ve enerji maliyetlerini minimize etmek için hedef programlama fonksiyonları yazılmıştır. Bu çalışmada aynı zamanda hat hızını da maksimum edecek amaç fonksiyonu hedef programlamaya eklenmiştir.

İkinci bölüm literatür taraması, üçüncü bölüm bulanık hedef programlama, bir sonraki bölüm demir çelik fabrikasındaki ısıl işlem bölümünün tanıtımı, kapasitesi ve işlem fonksiyonları hakkında bilgi verilecektir. Sonraki bölümde elde edilen sonuçların yorumlanması ve sonuç bölümü yer almıştır.

2. LİTERATÜR TARAMASI

Sürekli Tavlama Hattı için aşağıdaki çalışmalar incelenmiştir. Ayan, Singh, Prasadb, Chowdhurib(2012), yaptığı çalışmada yapay sinir ağı ve klasik matematiksel modeli birleştirerek sistemin daha etkili kontrolü için matematiksel modelde iyileştirmeler yapmıştır. Carvalho, Ong ve Guimaraes (2006), tavlama sürecini kontrol etmek için matematiksel modeli geliştirmişlerdir. Ming, Datai, Jiangang (2008), kullanılmayan kontrol model yerine yeni şerit kontrol modelini yapay zekâ metotlarını kullanarak geliştirmişlerdir. Christopher, Watanapong-se and Gaskey(1988), lineer kaudratik Gauss kontrol tekniklerini kullanarak araba sacı üretmek için kontrol tekniğini geliştirmişlerdir. Wang, Li, Hua ve Liu (2010), tavlama fırını için yeni süreç kontrol ve kendi-kendini uyarlama (self-adapting) modeli geliştirmişlerdir. Zhang, Shuai ve Tenga (2012), özyinelemeli Kernel temel bileşen analizi (Recursive Kernel principal component analysis - RKPCA) metodunu kullanarak güvenilir çalışma ortamı oluşturmak için süreci izleme, arızaları tespit etme ve giderme amacıyla çalışmalar yapmışlardır. Ghasem, Mahmoud, Vaghefi ve Lotfiani (2005), radyan tüplerin ömürlerini uzatmak için çalışmışlar. Shaoyuan, Chena ve Huang (2006), GGAP-RBF sinir ağını kullanarak kalite kontrolün dinamik modellemesi üzere çalışmışlardır. Yanagishima, Shimoyama, Suzuki, Sunami, Haga, Ida ve Irie (1981), Chiba Çelik Fabrikasında Sürekli Tavlama Hattı'nın karakteristik-lerini ve operasyonunu anlatmışlardır. Yoshitani ve Hasegawa (1998), vektör tipi değişken unutma faktörü ile özyinelemeli parametre kestirimi (recursive parameter estimation with a vector-type variable forgetting factor (REVVF)) metodunu kullanarak kendi kendini ayarlayabilen bir model geliştirmişlerdir. Fei, Yong-qin ve Shuren (2012), ısı transfer katsayılarını kendi kendine ayarlayabilen bir metot geliştirmişlerdir. Marlow (1996), kalite geçişleri sırasında tavlama fırını için matematiksel bir model tanımlamaktadır. Arumugam, Chandramohan ve Murthy (2011), yapay zekâ algoritmaları kullanarak birden fazla tavlama fırını olan işletmelerde fırın kullanım optimizasyonu için yeni model oluşturmuştur. Choa, Cassandras ve Kwon (2004), ileri ayrıştırma algoritması (forward decomposition algorithm) kullanarak birden fazla tavlama fırını olan işletmelerde fırın kullanım optimizasyonu için yeni model oluşturmuştur. Guo, Zhang, You, Chen ve Zhang (2009), parçacık sürü optimizasyonu algoritmaları (Particle swarm optimization) kullanarak fırın kontrol modeli geliştirmişlerdir. Bu modelde en uygun ısıtmak için gerekli gaz akışını PSO kullanarak optimize etmek istemişlerdir. Jeong, Yi, Kim ve Ha (1991), her zon için gaz sıcaklığını ölçerek kontrol metodu geliştirmişlerdir.

Hedef Programlama yaklaşımının birçok alan için pratik ve uygun olduğu yapılan çalışmalar ile anlaşılmaktadır. Chen, Chen ve Huang (2009) firma yöneticilerinin rekabet edebilirliğini arttırmak için büyük önem gösterdiği esnek üretim hücreleri için donanım alımı üzerine Tiwari'nin "Ağırlıklandırılmış Toplamsal Yaklaşımı" 'nı (Weighted Additive Approach) kullanarak toplam makina sayısı, toplam makinaların kapladığı alan, toplam satın alma maliyeti ve bu parti makinanın üretim miktarı arasındaki ilişkileri hesaplayan bir model geliştirmişlerdir. Selim, Araz ve Özkarahan (2008) tedarik zinciri sistemlerinde ortak üretim-dağıtım planlaması sorunu üzerinde yine Tiwari'nin "Ağırlıklandırılmış Toplamsal Yaklaşımı" (Weighted Additive Approach) yöntemini kullanarak uğraşmışlardır. Ertuğrul ve Güneş (2008) firmanın satış ve kar hedefleri için Narasihma yaklaşımını kullanan model geliştirerek bu sorunu çözmeye çalışmışlardır. Mékidiche, Mouslim and Sahed (2013) ana üretim planlaması için Hannan yaklaşımına yeni bir yaklaşım getirerek problem çözmeye çalışmışlardır. Güneş ve Umarosman (2005) bulanık hedef programlamayı kullanarak yeni bir yaklaşımla bulanık aritmetik ortalama üzerinde çalışmışlardır. Ertuğrul (2005) bir tekstil firmasının hazır giyim fabrikası ve ev tekstili grubuna önce doğrusal programlama sonra da hazır giyim fabrikasında satış ve kar hedefleri, ev tekstili grubunda ise satış hedefleri ile bulanık hedef programlama modeli uygulanarak iki modeli kıyaslamıştır. Erdin (2008) içme suyu üretimi için korunan doğal bir havzanın toprak örtüsüne müdahale edilmesi durumunda su kalitesi parametrelerindeki değişikliklerin, bulanık hedef programlama tekniklerinden "üçgensel üyelik fonksiyonlarıyla Chen yaklaşımı" kullanılarak değerlendirmesini yapmıştır. Liang (2006) tedarik zincirindeki üretim/nakliye planlama kararları için Zimmermann yaklaşımını kullanarak çözmeye çalışmıştır. Güneş ve Umarusman (2007) Yang, Ignizio ve Kim yaklaşımını kullanılarak, yerel yönetimdeki vergi uygulama problemi üzerinde çalışmalar yapmışlardır. Karaman ve Kale (2007) bir inşaat projesinin kalitesi, süresi ve maliyeti arasındaki ilişkinin nasıl olduğunu Tiwari Toplamsal Yaklaşımı yöntemini kullanılarak incelemiştir. Kağnıcıoğlu (2006) ana üretim planlaması için bulanık hedef programlama için Tiwari'nin toplamsal model yaklaşımını kullanarak çıkan değerleri hedef programlama ile kıyaslamıştır. Liang ve Cheng (2011) ana üretim planlaması için toplam maliyeti, toplam taşımayı, stok hacmini ve işçi sayısındaki değişimi dikkate alarak ürün çeşidi ve çoklu zaman aralıkları için yeni bir bulanık model geliştirmişlerdir. Özcan ve Toklu (2009) iki taraflı montaj hattı için Tiwari yaklaşımını kullanan ve iki montaj hattı dengelemesini yapan bir model önermişlerdir. Mekidiche ve Belmokaddem (2013) Yaghoobi ve Tamiz (2008)'in ana üretim planı için "Ağırlıklandırılmış Toplamsal Yaklaşım" metodu kullanarak geliştirmiş olduğu modele yeni bir formülasyon getirmişlerdir.

3. BULANIK HEDEF PROGRAMLAMA

Hedef programlama bir optimizasyon programıdır. Hedef programlama, çok kriterli karar analizi olarak bilinen çok kriterli karar vermenin bir dalı olan çok amaçlı optimizasyonun bir alt dalı olarak tanımlamaktadır. Hedef programlama aslında geliştirilmiş bir lineer programlamadır. Tek bir hedef yoktur. Birçok hedef vardır ve bu hedeflerin karar vericinin atadığı önem sırasına göre karşılanması gerekmektedir. Öncelik düzeyi yüksek olan hedeften başlayarak hedef düzeyi düşük olan hedefe doğru hedeften sapmaların minimize edilmesine çalışılır. Başarılı bir hedef programlama karar vericinin ihtiyaçlarını karşılamalıdır. Hedef programlamanın kuvvetli yönü olarak sadeliği ve kullanım kolaylığı olmasına rağmen zayıf yönü olarak Pareto yeterliliğini sağlayamamasını söyleyebiliriz.

Hedef programlama tekniği ilk olarak 1955 yılında Charnes, Cooper ve Ferguson tarafından kullanılmıştır (Charnes, Cooper, Ferguson (1955)). Fakat kullanımından 6 yıl sonra bu tekniğin adından ilk defa 1961 yılında Charnes ve Cooper 'ın makalesinde söz edilmiştir (Charnes ve Cooper (1961)). İlk mühendislik uygulaması ise aya ilk insanı indiren Apollo uzay kapsülünün aya konuşlandırılması işlemidir.

Bulanık Mantık ise birden fazla çözümü olan problem ve verilerle uğraşan matematiksel bir tekniktir. Burada sonsuz sayıda çözüm vardır. Klasik mantıkta olduğu gibi 0 veya 1 gibi kesin çözümler yoktur. 0 ile 1 arasında değişen sonsuz sayıda çözüm bulunabilmektedir. Bu nedenle insan mantığına en uygun programlama tekniğidir.

Bulanık küme mantığı 1960 yıllarında Kaliforniya Üniversitesinden Lütfi A. Zadeh tarafından geliştirilmiştir. 1987 yılına kadar kullanım alanı bulamamıştır. İlk olarak Matsushita firması çamaşır makinasının yüküne, kumaş karışımına ve deterjan miktarına göre yıkama periyodunu optimize eden bir uygulama yapmıştır. Bundan sonra ticari olarak çok geniş alanlarda uygulanmaya başlamıştır.

Bulanık hedef kümesi evrensel kümenin (seçenekler kümesinin) bir alt kümesi olan $\widetilde{W}$ kümesi ile veya $\mu_{\widetilde{W}}(\mathrm{x})$ üyelik fonksiyonu ile gösterilmiştir. Kısıtlayıcıların parametre değerleri ve/veya sağ taraf sabitleri bulanık olabilir. Bu durumda bulanık bir kısıtlayıcı evrensel kümede yer alan $\tilde{c}$ bulanık kümesi ile veya $\mu_{\tilde{c}}(\mathrm{x})$ üyelik fonksiyonu ile ifade edilmiştir. Bulanık karar kümesi ise evrensel kümede yer alan $\widetilde{D}$ bulanık kümesi ile veya $\mu_{\widetilde{D}}(\mathrm{x})$ üyelik fonksiyonu ile ifade edilmiştir[3].

$\widetilde{W}$ bulanık kümesi ile belirtilen hedeflere ulaşmak için $\tilde{c}$ bulanık kümesinin bütün kısıtlarının karşılanması gerekir. Bu durumda ise $\widetilde{W}$ ve $\tilde{c}$ bulanık kümelerinin kesişimlerini almamız gerekmektedir.

$$\tilde{D} = \tilde{W} \cap \tilde{c}$$

Bu kesişimi bulanık mantık yaklaşımı ile gösterirsek; n adet bulanık hedef ve m adet bulanık kısıtlayıcı olduğunda bulanık karar kümesi aşağıdaki gibi tanımlanır (Zimmermann,1991).

$\mu_{\tilde{D}}(x)$ = minimum $[\mu_{\tilde{W}}(x),\ \mu_{\tilde{c}}(x)]$;

$\forall\, x \in \mathcal{U}\ ;\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,m;$

$\mu_{\tilde{D}}(x) \in [0,1]$;

$\mu_{\tilde{W}}(x) \in [0,1]$;

$\mu_{\tilde{c}}(x) \in [0,1]$;

Özkan(2003), geleneksel karar verme problemini, karar verici, hedef, karar ölçütü, seçenekler, olaylar ve sonuç olmak üzere altı bileşenden oluştuğunu söylemiştir. Özkan (2003)'a göre bileşenlerden karar verici ve seçeneklerde bulanıklık olmadığı, hedef ve karar ölçütünün bulanıklık içerebileceği belirtilmiştir. Çünkü ulaşılmak istenen hedef düzeyi duruma göre bulanıklık gösterebilir. Mesela hedef 200 civarında denilebilir. Karar ölçütü içinde aynısı geçerlidir. Belirli bir limitin biraz altı gibi bulanık kelimeler ile ifade edilebilir.

Geleneksel doğrusal programlama problemi aşağıdaki denklem ile gösterilir.

Amaç fonksiyonu;

Maksimum Z = $\sum_{j}^{n} c_j x_j\ ;\ j = 1,2,\ldots,n$

Kısıtlayıcılar;

$\sum_{j=0}^{n} a_{ij} x_j \ \leq\ b_i\ ;\ i = 1,2,\ldots,m\ ;\quad x \geq 0$

Bu denklemlerde hedef programlama katsayıları (c_j), kısıtlayıcı katsayıları (a_{ij}) ve eşitsizlik sabitleri (b_i) değişmez. Bu denklemlerde bu kısıtlar altında en uygun değer bulunmaya çalışılır. Fakat bu katsayılar kesin ifade edilemeyebilir. Bir malın satış fiyatı gerektiğinde düşürülebilir veya arttırılabilir. Talep miktarı değişkenlik gösterebilir. Bazı yöntemler ile eşitsizlik katsayıları da değişebilir. Bu durumlarda bu katsayıların değerleri bulanık değerler ile gösterilmektedir.

Literatürde yaygın olarak kullanılan 4 tür bulanık doğrusal programlama vardır: bulanık kısıtlayıcılı doğrusal programlama,

Maksimum $(Z = c^T x)$

Kısıtlayıcılar

$(Ax)_i \lesssim b_i$

Bulanık amaç fonksiyonlu ve bulanık kısıtlayıcılı doğrusal programlama,

$\widetilde{\text{Maksimum}}$ $(Z = c^T x)$

Kısıtlayıcılar

$(Ax)_i \lesssim b_i$

Bulanık amaç katsayılı doğrusal programlama

Maksimum (Z = $\tilde{c}^T x$)

Kısıtlayıcılar

$(Ax)_i \leq b_i$

ve bulanık parametreli doğrusal programlama.

Maksimum (Z = $\tilde{c}^T x$)

Kısıtlayıcılar

$(\tilde{A}x)_i \leq \tilde{b}_i$

Özkan'a göre bulanık hedef programlama çözümlerinin çoğunda Zimmermann yaklaşımından esinlenilmiştir. Zimmermann 'a göre hedef bulanık amaç fonksiyonu probleme bulanık bir kısıt olarak katılabilir. Bu kısıtın boyutu karar verici tarafından belirlenmelidir. Bu yaklaşımın denklemsel gösterimi aşağıda gösterilmiştir.

$c^T x \gtrsim b_i$

Kısıtlayıcılar

$$\mu_i[(Bx)_i] = \begin{cases} 0;\ \text{eğer}\ (Bx)_i \geq b_i + d_i \\ \in [0,1]\ ;\ \text{eğer}\ b_i \leq (Bx)_i \leq b_i + d_i \\ 1;\ \text{eğer}\ (Bx)_i \leq b_i \end{cases}$$

$(Ax)_i \lesssim b_0$

Bu yaklaşıma göre problem;

$Bx \lesssim$ e ve $x \geq 0$ problemi haline gelir.

Bu problemin i. üyelik kümesi de aşağıdaki gibi olur. Bu üyelik fonksiyonuna göre eğer eşitsizlik tamamen karşılandı ise 1 değerini, hiç karşılanmadı ise 0 değerini alır. Karşılama derecesine göre 1'den 0'a doğru azalır. Bulanık hedefler için Zimmermann tipi üyelik fonksiyonları aşağıdaki gibi tanımlanır.

$(Ax)_i \cong b_i \; (i = 1, 2, \ldots, m_1) \Rightarrow$

$$\mu_i(x) = \begin{cases} 0 & ; \text{eğer} \;\; (Ax)_i \leq b_i - d_i \\ 1 - \frac{b_i - (Ax)_i}{d_i} & ; \text{eğer} \;\; b_i - d_i \leq (Ax)_i \leq b_i \\ 1 - \frac{(Ax)_i - b_i}{d_i} & ; \text{eğer} \;\; b_i \leq (Ax)_i \leq b_i + d_i \\ 1 & ; \text{eğer} \;\; (Ax)_i \geq b_i + d_i \end{cases}$$

$(Ax)_i \lessapprox b_i \; (i = m_1 + 1, \ldots, m_i) \Rightarrow$

$$\mu_i(x) = \begin{cases} 1 & ; \text{eğer} \;\; (Ax)_i \leq b_i \\ 1 - \frac{b_i - (Ax)_i}{d_i} & ; \text{eğer} \;\; b_i \leq (Ax)_i \leq b_i + d_i \\ 0; \text{eğer} \;\; (Ax)_i \geq b_i + d_i & \end{cases}$$

$(Ax)_i \gtrapprox b_i \; (i = m_2 + 1, \ldots, m_3) \Rightarrow$

$$\mu_i(x) = \begin{cases} 0 & ; \text{eğer} \;\; (Ax)_i \leq b_i - d_i \\ 1 - \frac{(Ax)_i - b_i}{d_i} & ; \text{eğer} \;\; b_i - d_i \leq (Ax)_i \leq b_i \\ 1 & ; \text{eğer} \;\; (Ax)_i \geq b_i \end{cases}$$

Bu fonksiyonlarda i. bulanık hedef için karar vericinin belirlediği eşik değerinin b_i olduğu farz edilmiştir.

4. SAHA ÇALIŞMASI

Bu bölümde Ereğli Demir Çelik Fabrikasına ait sürekli tavlama hattına ait veriler kullanılarak optimizasyon işlemi gerçekleştirilmiştir. Öncelikle, ısı transferi, kapasite, ısı balansı, kapasitenin tav çevrimine göre değişimi, tavlama sınırlılıkları, radyan tüp ısıtma bölüm kapasiteleri değerleri dikkate alınarak sistem modellemesi gerçekleştirilmiştir. Daha sonra sistemin amaç ve kısıtları belirlenerek hedef programlama işlemi için gerekli olan öncelik durumlarının belirlenmesi sağlanmıştır.

Isı Transferi

CAL hattı fırın bölümlerinde, malzeme ve ısıtma / soğutma donanımları arasındaki ısı transferi temel olarak ışıma ve/veya konveksiyon şeklinde gerçekleşir.

Bu mekanizmalarda gerçekleşen ısı transferi fonksiyonu:

Işıma fonksiyonu:

Q_r = A.B. E_s. H_r. $(T_2$ - $T_1)$ (1)

Konveksiyon fonksiyonu:

Q_c = A. H_c. $(T_2$ - $T_1)$ (2)

A = Isı transferi yüzeyi = W.L.2

W = Şerit genişliği

L = Isı transfer alanı net uzunluğu

E_s = Işıma yayım kabiliyeti katsayısı (emissivity) (Işıma yayılım Katsayısı)

B = Kara cisim katsayısı

H_r, H_c = Isı transferi katsayıları

2 çarpanı şeridin 2 yüzeyi için kullanılmıştır

Kapasite

P = g.W. T.V.60 (3)

W = Malzeme şerit genişliği

T = Malzeme şerit kalınlığı

G = Özgül ağırlık

V = Proses Hızı

P = Proses kapasitesi

Bu kapasite için gerekli olan ısı miktarı;

Q_s = C_p. $(T_2$ - $T_1)$.P (4)

C_p = Özgül Isı

Isı Balansı

Isı eşitlik denklemlerinden;

Alınan ısı = Verilen ısı

Q_s =Q_r+Q_c (Bu denklemde % 10'dan az olan fırın kayıpları ihmal edilmiştir.)

(1), (2), (3) ve (4) numaralı denklemleri kullanarak;

C_p.$(T_2$ - $T_1)$.g.W.t.V.60=A.[H_r.B. E_s+H_c]

C_p.g.W.t.V.60 =W.L. [H_r.B. E_s+H_c]

g.W.t.V.60 $= \frac{\text{W.L.}[H_r.\text{B}.E_s+H_c]}{C_p}$

P= f(W.L)

Isıtma ve soğutma işlemlerinde ısı transfer mekanizmalarının değişmediği varsayılırsa; kapasitesinin şerit genişliği ve ısı transfer alanı net uzunluğunun (fırın uzunluğu) fonksiyonu olduğu görülmektedir.

Kapasite bu iki değişken ile orantılıdır.

C_p.g.t.V.60=W.L.[H_r.B. E_s+H_c]

t.V $= \frac{\text{L}.[H_r.\text{B}.E_s+H_c]}{C_p.g.60}$

"Şerit kalınlığı x Hız" değeri ısıtma ve soğutma işlemleri ile fırın uzunluğuna bağlı olarak sabittir.

Kapasitenin tav çevrimine göre değişimi

Q_r + Q_c toplamı fırın kuruluş standartlarına göre sabit kabul edilirse;

$Q_s = C_p .(T_2 - T_1)$.P= $Q_r + Q_c$ = sabit

Tavlama Sınırlılıkları:

Ürün Kalitesi	Şerit Sıcaklığı (°C)	Zaman (Sn.)
CQ	720 ± 30	≥ 40
DQ-1	770 ± 30	≥ 40
DQ-2	780 ± 30	≥ 40
DDQ	800 ± 30	≥ 40
EDDQ	830 ± 30	≥ 40

Hat hızı maksimum 320 m/d. olacaktır.

Radyan Tüp Isıtma Bölümü Kapasitesi

Bölüm	Bölge	Tüp Adedi	Teorik Gerçek Kapasite Kcal/H	
			Tüp	Zon
RTH1	1	59	85 000	4 900 000
RTH1	2	58	85 000	4 975 000
RHT1	3	69	85 000	5 649 000
RHT1	4	70	85 000	5 621 000
RTH2	5	29	120 000	3 460 000
RTH2	6	30	120 000	3 626 000
RTH2	7	30	120 000	3 260 000
RTH2	8	30	120 000	3 260 000
TOTAL	8	375	-	34430 000

Amacımız yukarıdaki verileri kullanarak hattın ısı kapasitesi sabit olduğu için kalite geçişlerindeki harcanacak olan enerjiyi minimum yapacak olan bulanık hedef programlama denklemi gerçekleştirmektir. Burada kısıtlar şeridin kalite bazına olan tavlama sıcaklıkları, şeridin tavlama bölgesindeki kalma zamanı (40 saniyeden büyük olmalı) ve ısıtıcıların ısıtma değerleridir. Bu denklemde;

Amaçlar:

1. Üretimi maksimize etmektir. (Kalite geçişlerinde her genişlik ve kalınlık değeri için en uygun ısı ve hız geçişlerini bulmak.)

Kısıtlamalar:

1. Kalitelerin belirli sıcaklıklar aralıklar arasında çekilmesini sağlamak.
2. Isıtıcıların verebileceği maksimum kalori miktarı
3. Şeridin alabileceği genişlik ve kalınlık sınırlamaları
4. Fırının alabileceği maksimum şerit uzunluğu
5. Hattın maksimum ve minimum hız miktarı (Üretim hızı şeridin 40 m/dakika içinde fırın bölgesinden geçilmesini sağlamalı.)

Bu işlemler için en fazla üretilen 6 adet kalınlık ve genişlik aralığı kullanılmıştır.

Kalınlık Aralığı	Genişlik Aralığı
0.40 - 0.50	1100 – 1200
0.51 - 0.60	1200 – 1300
0.71 - 0.80	1200 – 1300
0.81 - 0.90	1200 – 1300
0.91 - 1.00	1200 – 1300
1.01 - 1.20	1000 – 1100

Gerekli ısı miktarı = C_p * g * w * t * V * 60 * $(T_2 - T_1)$ * L

C_p = özgül ısı

g = özgül ağırlık

w = şerit genişliği

t = şerit kalınlığı

v = şerit hızı

T_n = sıcaklık

L = fırın uzunluğu

İşlem kısıtları:

Kalınlık sınırlamaları:

$0.40 \text{ mm} \leq t \leq 2.00 \text{ mm}$

Genişlik sınırlamaları:

$625 \text{ mm} \leq w \leq 1525 \text{ mm}$

Hız sınırlamaları:

$30 \text{ m/d.} \leq v \leq 320 \text{ m/d.}$

Kalite sıcaklıkları sınırlamaları:

CQ kalite için; 720 ± 30 °C

DQ-1 kalite için; 770 ± 30 °C

DQ-2 kalite için; 780 ± 30 °C

DDQ kalite için; 800 ± 30 °C

EDDQ kalite için; 850 ± 20 °C

Fırın uzunluğu = RTH1 uzunluk + RTH2 uzunluk

L = 292 + 278 = 570 metre

Toplam ısıtma kapasitesi fırın boyunca (375 adet radyan tüp için)

Değişkenler:

CQ kalite için; DDQ kalite için;

v_1 Şerit hızı,

t_1 Şerit kalınlığı,

w_1Şerit genişliği,

x_1 Şerit sıcaklığı

DQ-1 kalite için;

v_2 Şerit hızı,

t_2 Şerit kalınlığı,

w_2Şerit genişliği,

x_2 Şerit sıcaklığı

DQ-2kalite için;

v_3 Şerit hızı,

t_3 Şerit kalınlığı,

w_3Şerit genişliği,

x_3 Şerit sıcaklığı

v_4 Şerit hızı,

t_4 Şerit kalınlığı,

w_4 Şerit genişliği,

x_4 Şerit sıcaklığı

EDDQ kalite için;

v_5 Şerit hızı,

t_5 Şerit kalınlığı,

w_5 Şerit genişliği,

x_5 Şerit sıcaklığı

<u>Sabitler:</u>

Özgül ağırlık: 7,80 10^3 kg/m^3

Özgül ısı: 0,1170 kilo Kalori / kg °C

Fırın Isı Kapasitesi = 3440000 Cal*m / h

Sabit C = C_p * g * 60 * L

<u>Denklemlerin ve Kısıtların oluşturulması:</u>

0,40 mm ≤ t_1, t_2, t_3, t_4, t_5 ≤ 2,00 mm

625 mm ≤ w_1, w_2, w_3, w_4, w_5 ≤ 1525 mm

1800m/saat ≤ v_1, v_2, v_3, v_4, v_5 ≤ 19200 m/saat

690 °C ≤ x_1 ≤ 750 °C

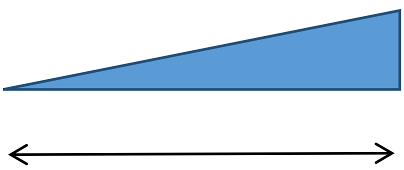
740 °C ≤ x_2 ≤ 800 °C

750 °C ≤ x_3 ≤ 810 °C

770 °C ≤ x_4 ≤ 830 °C

830 °C ≤ x_5 ≤ 870 °C

<u>Bulanık Hedef Programlama Denklemlerinin ve Kısıtların Oluşturulması</u>



1800 $\boldsymbol{d_i}$ $\boldsymbol{b_i = 19200}$

Tiwari, Dharmar ve Rao'nun Toplamsal Model yaklaşımına uygun üyelik fonksiyonumuz;

$$\mu_i\,(x) = \left\{1 - \frac{(b_i - v_i(x))}{d_i};\right. \qquad b_i - d_i \le v_i(x) \le b_i$$

$$1 - \frac{(19200 - v_1)}{16800} = \mu_1 \text{ (m/h)}$$

$$1 - \frac{(19200 - v_2)}{16800} = \mu_2 \text{ (m/h)}$$

$$1 - \frac{(19200 - v_3)}{16800} = \mu_3 \text{ (m/h)}$$

$$1 - \frac{(19200 - v_4)}{16800} = \mu_4 \text{ (m/h)}$$

$$1 - \frac{(19200 - v_5)}{16800} = \mu_5 \text{ (m/h)}$$

$C * w_1 * t_1 * v_1 * (x_1 - x_0) \le 3440000$

$C * w_2 * t_2 * v_2 * (x_2 - x_1) \le 3440000$

$C * w_3 * t_3 * v_3 * (x_3 - x_2) \le 3440000$

$C * w_4 * t_4 * v_4 * (x_4 - x_3) \le 3440000$

$C * w_5 * t_5 * v_5 * (x_5 - x_4) \le 3440000$

$0{,}40\ 10^{-3}\text{ m} \le t_1 \le 2{,}00\ 10^{-3}\text{m}$

$0{,}40\ 10^{-3}\text{m} \le t_2 \le 2{,}00\ 10^{-3}\text{m}$

$0{,}40\ 10^{-3}\text{m} \le t_3 \le 2{,}00\ 10^{-3}\text{m}$

$0{,}40\ 10^{-3}\text{m} \le t_4 \le 2{,}00\ 10^{-3}\text{m}$

$0{,}40\ 10^{-3}\text{m} \le t_5 \le 2{,}00\ 10^{-3}\text{m}$

$625\ 10^{-3}\text{m} \le w_1 \le 1525\ 10^{-3}\text{m}$

$625\ 10^{-3}\text{m} \le w_2 \le 1525\ 10^{-3}\text{m}$

$625\ 10^{-3}\text{m} \le w_3 \le 1525\ 10^{-3}\text{m}$

$625\ 10^{-3}\text{m} \le w_4 \le 1525\ 10^{-3}\text{m}$

$625\ 10^{-3}\text{m} \le w_5 \le 1525\ 10^{-3}\text{m}$

$1800 \le v_1 \le 19200$ m/h

$1800 \le v_2 \le 19200$ m/h

$1800 \le v_3 \le 19200$ m/h

$1800 \le v_4 \le 19200$ m/h

$1800 \le v_5 \le 19200$ m/h

690 °C ≤ x_1 ≤ 750 °C

740 °C ≤ x_2 ≤ 800 °C

750 °C ≤ x_3 ≤ 810 °C

770 °C ≤ x_4 ≤ 830 °C

830 °C ≤ x_5 ≤ 870 °C

$\mu_1 \ge 0$, $\mu_2 \ge 0$, $\mu_3 \ge 0$, $\mu_4 \ge 0$, $\mu_5 \ge 0$

Ağırlıkların bulunması için yıllık üretim miktarı oranları temel alınmıştır. Buna göre CQ kalite için değişkenlerin ağırlıkları %60, DQ-1 kalite için değişkenlerin ağırlıkları %11, DQ-2 kalite için değişkenlerin ağırlıkları %11, DDQ kalite için değişkenlerin ağırlıkları %15 ve EDDQ kalite için değişkenlerin ağırlıkları %3 olarak alınmıştır.

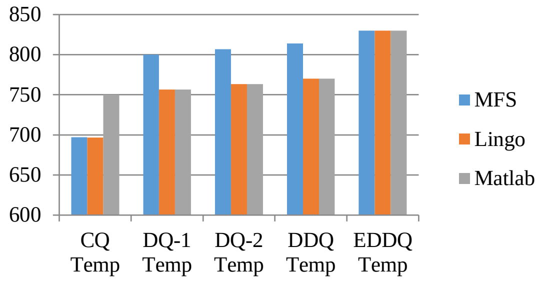
Buna göre oluşturulan maksimizasyon denklemi:

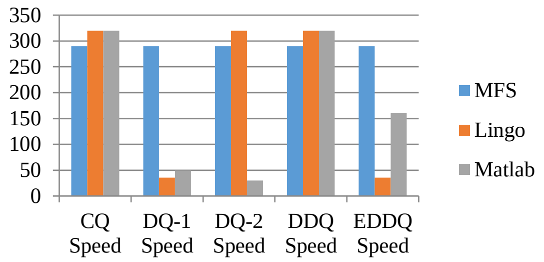
$$\textbf{Maksimum}\ \sum \{\mathbf{0.6} * \mu_1 + \mathbf{0.11} * \mu_2 + \mathbf{0.11} * \mu_3 + \mathbf{0.15} * \mu_4 + \mathbf{0.03} * \mu_5\}$$

5. SONUÇLAR

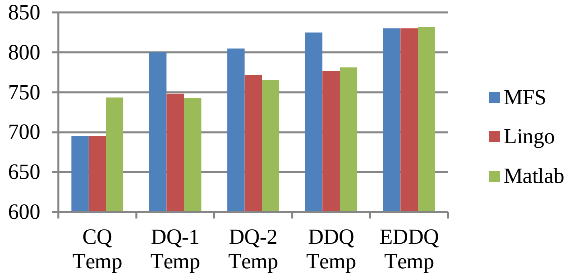
Bulanık hedef programlama denklemleri ve kısıtları, üretim hatlarında en çok üretilen ebatlar için Lingo, Microsoft Solver Foundation ve Matlab Fmincon yazılımları ile denenmiştir. Programların çalıştırılması sonucu üretilen sonuçlar grafik olarak aşağıdaki gibi bulunmuştur.

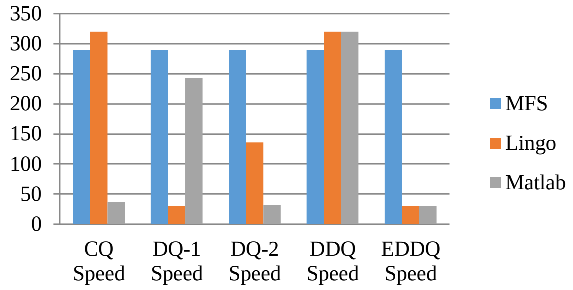
Grafik 1. 0.45 mm kalınlık ve 1150 mm genişlik için en uygun hız ve sıcaklık grafiği



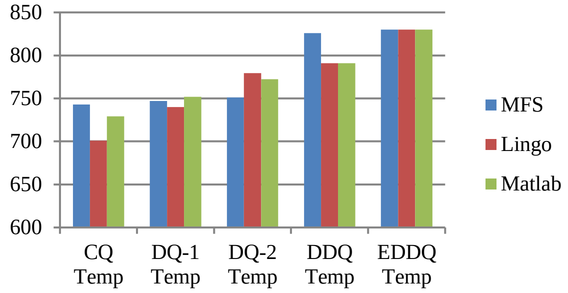


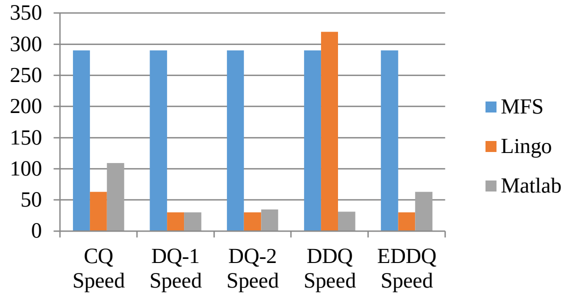
Grafik 2. 0.55mm kalınlık ve 1250 mm genişlik için en uygun hız ve sıcaklık grafiği



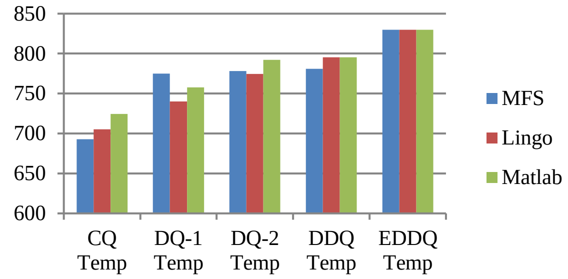


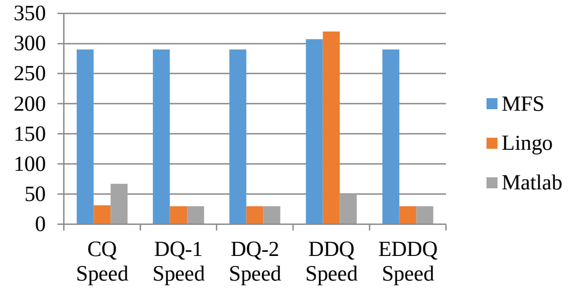
Grafik 3. 0.75 mm kalınlık ve 1250 mm genişlik için en uygun hız ve sıcaklık grafiği



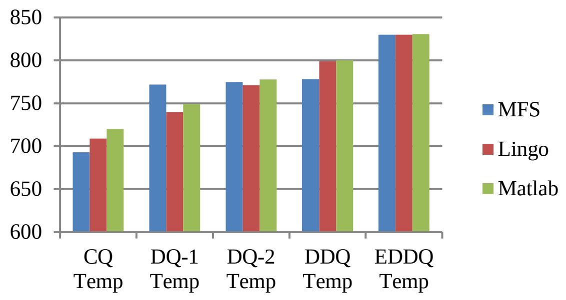


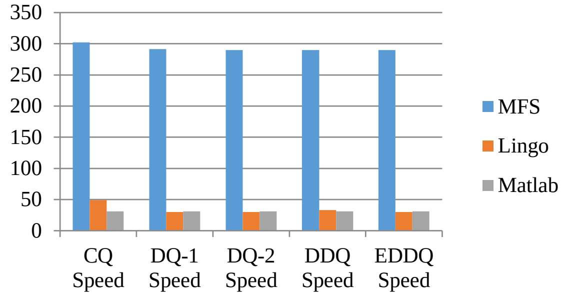
Grafik 4. 0.85 mm kalınlık ve 1250 mm genişlik için en uygun hız ve sıcaklık grafiği





Grafik 5. 0.95 mm kalınlık ve 1250 mm genişlik için en uygun hız ve sıcaklık grafiği





6. YORUM VE GELECEK İÇİN ÖNERİLER

Yukarıdaki grafiklerden görüleceği üzere ince en yüksek hız performansını bulmada Microsoft firmasının Solver Foundation 3.0 çerçevesinin içinde tanımlandığı "HybridLocalSearchDirective" en iyileme algoritması sağlamıştır. Lingo yazılımı için NLP sınıfı en iyileme sınıfı ve matlab için yukarıdaki hedef formülde değişiklik yapılarak (1- max V(λ)) fmincon (Find minimum of constrained nonlinear multivariable function) fonksiyonu kullanılmıştır. Fakat genelde Lingo ve Matlab'ın en iyileme fonksiyonlarının ürettiği sonuçlar pek tatmin edici olmamıştır. Fakat yapılan çalışmalarda yine de 290 m/dk. Hıza ulaşmak mümkün olabilmiştir. Maksimum üretim kapasitesine ulaşmak, dolayısıyla üretim maliyetlerini düşürmek için çalışmalara sezgisel algoritmaları da dâhil ederek çalışmalara devam etmek gerekmektedir. Bu sayede enerji verimliliği de sağlanmış olacaktır çünkü enerji verimliliği sağlamanın bir yolu da sıcaklık geçişlerinin az olmasında yatmaktadır. Gelecekte hızı maksimum hız olan 320 (m/h) 'ye çıkarmak için çalışmalara devam edilecektir.

References

Arumugam, M. S., Chandramohan , A., Murthy G. R.(2011). On the optimal control of steel annealing processes via various versions of genetic and particle swarm optimization algorithms. International Journal of Bio-Inspired Computation, 12, 371-392

Carvalho, S.R., Ong, T.H. ve Guimaraes, G.(2006). A mathematical and computational model of furnaces for continuous steel strip processing. Journal of Materials Processing Technology , 178, 379 – 387

Charnes A., Cooper, W.W. ve Ferguson,R. (1955). Optimal estimation of executive compensation by linear programming. Journal of Management Science, 1, 138-151.

Chen, Y., Chen, L. and Chien-Yu Huang. (2009). Fuzzy Goal Programming Approach to Solve the Equipment - Purchasing Problem of an FMC. International Journal of Industrial Engineering, 270-281.

Choa, Y.C. , Cassandrasb, C. G. ve Kwon W. H. (2004). Optimal control for steel annealing processes as hybrid systems. Journal of Control Engineering Practice, 12, 1319–1328

Dini, G. , Vaghefi, S.M.M. ve Lotfiani,M. (2008). Computational and experimental failure analysis of continuous-annealing furnace radiant tubes. Journal of Engineering Failure Analysis, 15, 445–457.

Erdin, Dr. C. (2008). Bulanık Hedef Programlama ve Su havzasında bir uygulama. Journal of İstanbul University Social Sciences, 89-108

Erkan, K. ve Kale, S. (2007) Bulanık Hedef Programlama Yöntemi ile Süre-Maliyet-Kalite Eniyilemesi. Construction Management Congress, 109-118

Ertuğrul, İ. (2005). Bulanık Hedef Programlama Ve Bir Tekstil Firmasında Uygulama Örneği. Journal of Eskişehir Osmangazi University Social Sciences, 45-78

Ertuğrul, İ. ve Güneş, M. (2007). Fuzzy Goal Programming and an Application of Production Process. Analysis and Design of Intelligent Systems using Soft Computing Techniques Advances in Soft Computing, 41, 649-659.

Goswami, A., A., Singh, R.K., Prasad, A., Chowdhuri, Dr.S. (2012). Mathematical model based control system for Silicon Steel Mill of Rourkela Steel Plant. Journal of Procedia Technology, 4, 867-872

Guo C. , Zhang, Y. , You, X., Chen X. and Zhang, Y.(2009). Optimal Control of Continuous Annealing Process Using PSO. Int. Conference on Automation and Logistics, 602-606

Güneş, M. ve Umarosman, N. (2005). Fuzzy Goal Programming Approach on Computation of the Fuzzy Arithmetic Mean. Mathematical and Computational Applications, 211-220

Güneş, M. ve Umarusman N. (2007). Bir Karar Destek Aracı Bulanık Hedef Programlama ve Yerel Yönetimlerde Vergi Optimizasyonu Uygulaması. Review of Social, Economic & Business Studies, 242-255

Hans, J. Z. (1991). Fuzzy Set Theory and its Application. Kluwer Academic Publisher, Boston.

Jeong, S. H., Yi, J.J., Kim, J.K. ve Ha, M.Y. (1991). Computer Modeling of the Continuous Annealing Furnace. Journal of KSME, 5, 16-21

Kağnıcıoğlu, C.H. (2006). Hedef Programlama ve Bulanık Hedef Programlama Arasındaki ilişki. Journal of Gazi University Economics and Administrative Sciences Faculty, 17-38

Kelly, C. D. , Watanapongse, D. ve Gaskey, K. M.(1998). Application of Modern Control to a Continuous Anneal Line. IEEE Control Systems Magazine, 32-37.

Lia, S., Qing Chena, Q. ve Huang, G.(2006). Dynamic temperature modeling of continuous annealing furnace using GGAP-RBF neural network. Journal of Neurocomputing, 69, 523–536

Liang, T. (2006). Applying fuzzy goal programming to production / transportation planning decisions in a supply chain. International Journal of Systems Science, 293-304

Liang, T. ve Cheng, H. (2011). Multi-Objective Aggregate Production Planning Decisions Using Two-Phase Fuzzy Goal Programming Method. Journal Of Industrial And Management Optimization, 7, 365-385.

Marlow D.O. (1996). Modelling direct-fired annealing furnaces fortransient operations. Journal of Application Mathematics Modelling, 20, 34-40

Mekidiche, M. ve Belmokaddem, M. (2013). Weighted Additive Fuzzy Goal Programming Approach to Aggregate Production Planning. I.J. Intelligent Systems and Applications, 4, 20-29.

Mekidiche, M., Hocine, M. ve Abdelkader S. (2013). Application of tolerance approach to fuzzy goal programming to aggregate production planning. Int. J. Mathematics in Operational Research, 5, 183-204.

Özcan, U. ve and Toklu, B. (2009). Multiple-criteria decision-making in two-sided assembly line balancing: A goal programming and a fuzzy goal programming models. Journal of Computers & Operations Research, 36, 1955 - 1965

Özkan, M. M. (2003). Bulanık Hedef Programlama. Ekin Kitabevi, İstanbul.

Selim, H. , Araz, C. ve Özkarahan İ. (2008). Collaborative production–distribution planning in supply chain: A fuzzy goal programming approach. Transportation Research Part E: Logistics and Transportation Review, 396-419.

Sert, Ü. (2015).Haddeleme Hakkında. http://www.arsivbelge.com/yaz.php?sc=2418

Wan F., Wang Y. ve Qin S.(2012). Modeling of Strip Heating Process in Vertical Continuous Anealing Furnace, Journal Of Iron And Steel Research, 19, 29-36

Wen-Ie, W., Jian-ping, L. , Fu-an, H. ve Xiang-hua L.(2010). Application of Self-Learning to Heating Process Control of Simulated Continuous Anneal. Journal Of Iron And Steel Research, 17, 27-31

Yaghoobi, M.A., Jons, D.F., Tamiz. (2008). Weighted additive models for solving fuzzy goal programming problems. Int. Journal of Intelligent Systems and Applications. Asia Pacific Journal of Operation Research , 25, 715-730

Yanagishima,F., Shimoyama, Y., Suzuki,M., Sunami, H., Haga, T., Ida, Y. ve Irie T.(1981). Characteristics and Operation of Multipurpose Continuous Annealing Line at Chiba Works. Kawasaki Steel Technical Report, 2, 1-13

Yoshitani, N. ve Hasegawa A. (1998). Model-Based Control of Strip Temperature for the Heating Furnace in Continuous Annealing, IEEE Transactions on Control Systems Technology, 6, 146-156

Zhang, Y. , Li, S. ve Tenga, Y. (2012). Dynamic processes monitoring using recursive kernel principal component analysis. Journal of Chemical Engineering Science , 72, 78–86

Zhou, M., Yu D. ve Zhou J. (2008). A New Strip Temperature Control Method for the Heating Section of Continuous Annealing Line. Cybernetics and Intelligent Systems 2008 IEEE Conference, 861 – 864