



POLİTEKNİK DERGİSİ

JOURNAL of POLYTECHNIC

ISSN: 1302-0900 (PRINT), ISSN: 2147-9429 (ONLINE)

URL: <http://dergipark.org.tr/politeknik>



Efficiency analysis of science and technology parks using data envelopment analysis: Evidence from Turkey

Teknoloji geliştirme bölgelerinin veri zarflama analizi kullanarak etkinliklerinin analizi: Türkiye örneği

Yazar(lar) (Author(s)): Necla ARSLAN¹, Önder BELGİN²

ORCID¹: 0000-0003-0963-6630

ORCID²: 0000-0001-6702-2608

Bu makaleye şu şekilde atıfta bulunabilirsiniz (To cite to this article): Arslan N. and Belgin Ö., “Efficiency analysis of science and technology parks using data envelopment analysis: Evidence from Turkey”, *Politeknik Dergisi*, 24(4): 1667-1674, (2021).

Erişim linki (To link to this article): <http://dergipark.org.tr/politeknik/archive>

DOI: 10.2339/politeknik.649833

Efficiency Analysis of Science and Technology Parks using Data Envelopment Analysis: Evidence from Turkey

Highlights

- ❖ Efficiency of 22 science and technology parks in Turkey is evaluated using data envelopment analysis
- ❖ Six different data envelopment model is used to examine the strong and weak areas of science and technology parks
- ❖ Efficiency scores obtained by the models are clustered using K-means method.

Graphical Abstract

Efficiency of 22 science and technology parks in Turkey is evaluated by six different data envelopment analysis models. Then, the science and technology parks are clustered depending on the obtained efficiency scores.

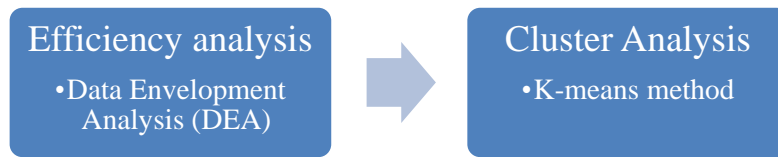


Figure. Methodological frame of the study

Aim

This study addresses the efficiency analysis of STPs in Turkey using Data Envelopment Analysis (DEA).

Design & Methodology

Firstly, six different Data Envelopment Analysis (DEA) models are employed and after that the STPs are clustered using K-means method.

Originality

This study is first on the efficiency of the STPs in Turkey employing additional analysis

Findings

Five of 22 STPs are found to be efficient and STPs exhibits lower performance in the efficiency of revenue and patents. Also, STPs can be clustered as Marketers, Researchers and Low-performers.

Conclusion

According to the analysis, there is a problem in the commercialization of the R&D projects. In addition to this, the gather patents from the output of the projects is problematic. It seems the innovativeness of the R&D projects needs to be increased.

Declaration of Ethical Standards

The author(s) of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

Efficiency Analysis of Science and Technology Parks Using Data Envelopment Analysis: Evidence from Turkey

Araştırma Makalesi / Research Article

Necla ARSLAN*, Önder BELGİN

Sanayi ve Teknoloji Bakanlığı, Türkiye

(Geliş/Received : 22.11.2019 ; Kabul/Accepted :02.06.2020 ; Erken Görünüm/Early View : 04.06.2020)

ABSTRACT

Research and Development (R&D) and innovation have a significant impact on the competitiveness of countries. Science and Technology Parks (STPs) are an important component of R&D and innovation ecosystems of countries and they aim to increase the university-industry collaboration. This study addresses the efficiency analysis of STPs in Turkey using Data Envelopment Analysis (DEA). For this purpose, an input-oriented DEA model is used to obtain efficiency scores of STPs and 5 of 22 STPs are found to be efficient. After that, to examine the strong and weak areas of STPs six additional Data Envelopment Analysis (DEA) models are considered. According to these models, STPs exhibits lower performance in the efficiency of revenue and patents. Finally, STPs are clustered based on efficiency scores as Marketers, Researchers and Low-performers using K-means clustering and we made suggestions for each cluster. The motivation of this study is contributing to policies for increasing the performance and the impact of the STPs in Turkey.

Keywords: Science and technology parks, data envelopment analysis, efficiency, cluster analysis.

Teknoloji Geliştirme Bölgelerinin Veri Zarflama Analizi Kullanarak Etkinliklerinin Analizi: Türkiye Örneği

ÖZ

Ar-Ge ve inovasyon ülkelerin rekabet edebilirliği üzerinde önemli bir etkiye sahiptir. Teknoloji Geliştirme Bölgeleri (TGB) ülkelerin Ar-Ge ve inovasyon ekosistemlerinin önemli bir bileşenidir ve üniversite sanayi işbirliğini artırmayı hedeflemektedirler. Bu çalışma Veri zarflama analizini (VZA) kullanarak Türkiye de bulunan TGB'lerin etkinlik analizi üzerinedir. Bu amaçla, girdi yönelimli VZA modeli TGB'lerin etkinlik skorlarını elde etmek için kullanılmıştır ve 22 TGB'nin 5'inin etkin olduğu bulunmuştur. Sonrasında, TGB'lerin güçlü ve zayıf olduğu alanları incelemek üzere altı ek veri zarflama analizi modeli dikkate alınmıştır. Bu modellere göre, TGB'ler patent ve gelir etkinliğinde düşük bir performans sergilemektedir. Son olarak, TGB'ler K-ortalamalar kümeleme yöntemi kullanarak Ticarileşmede Başarılı, Araştırmacı ve Düşük Performanslı olarak etkinlik puanlarına göre kümelenebilir ve her bir küme için önerilerde bulunulmuştur. Bu çalışmanın motivasyonu TGB'lerin Türkiye'deki performansını ve etkisini artırmaya yönelik politikalara katkıda bulunmaktır.

Anahtar kelimeler: Teknoloji geliştirme bölgeleri, veri zarflama analizi, etkinlik, kümeleme analizi.

1. INTRODUCTION

In today's competitive environment, science and technology are inevitable for the economic development of the countries. Especially developed countries transform the information and technology into social and economic contribution by developing university industry cooperation. Science and Technology Parks (STPs) plays an important role to achieve this cooperation effectively by R&D and innovation activities and producing high value-added products.

Today, there are more than 400 science and technology parks worldwide and this number continues to increase. With more than 150 technology parks, USA has the most number of STPs. In Turkey, STPs entered into legislation with the Technology Development Zones Law No. 4691 published in 2001. In this respect, STPs are used in the literature in the sense of *technopark* and two important features come to the forefront. The first one is related to high-technology production and secondly, they aim to develop companies by hosting their near universities or research centres for the commercialization of high technology outputs.

There are 60 active STPs in Turkey and these STPs hold 5.216 firms in their bodies. Nearly 50.000 R&D

*Sorumlu Yazar (Corresponding Author)
e-posta : necla.arslan@sanayi.gov.tr

personnel are employed in STP firms and 41% of the firms work in the field of software. Total export value of all the firms in STPs has reached to 3,6 billion USD.

In the literature it is possible to encounter many studies on science and technology parks. These studies can be grouped as empirical studies, case studies, theoretical or conceptual publications, literature reviews, and publications related to evaluation of STPs [1]. Since our study is about the efficiency of the STPs, we give the literature on this limited area. Hu et al. [2] analysed the efficiency of industrial parks in Taiwan with the Data Envelopment Analysis (DEA) approach. They studied with five input variables, including the number of firms, firm area, capital stock, utility maintenance fee, and number of employees and one output, operating value. In order to test the difference between the efficiency scores of four areas in Taiwan, they used Mann-Whitney test. This result showed that the north area has the best ranking of industrial park efficiency. Then, the researchers used the tobit regression, which is a censored regression designed to estimate the linear relationships between variables, to analyse the effects of environmental variables on the efficiency scores. They found out that the industrial zone with higher local unemployment rates is more efficient. Hu et al. [3] examined the efficiency of fifty-three Science and Technology Industrial Parks (STIPs) in China from 2004 to 2006. They applied four-stage (DEA) in order to examine the environmental factors effecting the efficiency of the STIPs. The researchers used number of firms, employees in an STIP, the percentage of employees having graduated from a university to total employees, R&D expenditures and the percentage of science and technology personnel to total employees as input variables; technical revenue, product sales revenue and commodity sales revenue as output variables. Then, they applied tobit regression to analyse the environmental effects. They find out the STIPs located in the east areas reveal more efficiency than those in the central and west area.

It is of great importance to evaluate the performances of the STPs in order to evaluate the development of them and to modify the policies in this area. In this study, performance of the STPs are evaluated using DEA firstly. After that, strong and weak areas of STPs are evaluated based on six different DEA models. Finally, STPs are clustered based on efficiency scores obtained previous DEA models. To the best of our knowledge, there isn't any study on the efficiency of the STPs in Turkey employing additional analysis and this study aims to fill the gap in this area.

This paper is organised as follows. After this introduction, the second section gives methodology including Data Envelopment Analysis and Cluster Analysis. Section three presents the results of the empirical study. The article ends with conclusions and future research directions.

2. METHODOLOGY

2.1. Data Envelopment Analysis

DEA is a nonparametric method which was first introduced by Farrel [4] and was modified by Charnes-Cooper and Rhodes [5]. Farrel [4] put forward a basic theory based on a single input and output. Then, Charnes-Cooper and Rhodes [5] used linear programming to compare multiple inputs and outputs to compare activities between decision-making units. They referred to this method as CCR. Another type of DEA model is BCC (Banker, Charnes and Cooper) by Bankers et al. [6]. CCR model considers constant returns of scale and this model gives the Technical Efficiency (TE) of each Decision-Making Units (DMUs) and BCC model considers variable returns of scale. This model gives Pure Technical Efficiency (PTE), which is a measure of efficiency without Scale Efficiency (SE) of each DMUs. In this study CCR model is used to obtain the efficiency of STPs.

In its most commonly used form, DEA is used to compute a score, which defines the relative efficiency of a particular DMU versus all other DMUs observed in the sample [7]. DEA allows measurement of efficiency from multiple inputs and multiple outputs within multiple DMUs [8]. DEA uses a linear programming approach to identify the efficient DMUs, those units that make the most efficient use of inputs to produce outputs. The efficiency units consist of a frontier among all DMUs [9].

Accordingly, the mathematical equation to find the maximum efficiency of DMUs using weighted input-output efficiency measure with constant returns of scale (CCR model) can be expressed as follows:

$$\max E_j = \sum_{r=1}^s u_r y_{r0} \quad (1)$$

Subject to

$$\sum_{i=1}^m v_i x_{i0} = 1 \quad (2)$$

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0 \quad (3)$$

$$j = 1, 2, \dots, n \quad (3)$$

$$u_r \geq 0 \quad r = 1, 2, \dots, s \quad (4)$$

$$v_i \geq 0 \quad i = 1, 2, \dots, m \quad (5)$$

where E_j is the relative efficiency of DMU_0 , u_r and v_i are weights assigned to output r and input i , respectively and y_{r0} and x_{i0} are the input and output data for DMU_0 . n , s and m state the number of DMUs, outputs and inputs successively. If E_j is equal to 1 the DMU is efficient, otherwise DMU is inefficient.

DEA can be implemented using either an input or an output orientation. In the input orientation approach, the objective is to estimate the degree of potential input savings for a given realized output level of the unit [10]. However, under the output-orientation, is measured the extent to which output may have been expanded for the level of inputs used by the unit [11]. We employed an input-oriented model for the corresponding analysis.

DEA models have been developing for last 50 years and for detailed information on the applications of DEA, Cook and Seiford [12] and Liu et al. [13,14] can be read.

2.2. Cluster Analysis

Cluster analysis is the identification of groups of observations that are cohesive and separated from other groups [15]. The main aim of clustering analysis is to make the in-group similarity maximum and to make the clustering similarity between the clusters. Clustering, just as the dimensionality reduction methods, can be used for two purposes. First, it can be used for data exploration to understand the structure of data, and second, to map data to a new space where supervised learning is easier [16].

Clustering is useful in several exploratory pattern analysis, grouping, decision-making and machine learning situations including data mining, document retrieval, image segmentation and pattern classification [17].

Different clustering methods may generate different clusters on the same data set. The partitioning is not performed by humans, but by the clustering algorithm. Hence, clustering is useful in that it can lead to the discovery of previously unknown groups within the data [18]. The two main ways to cluster data—make the partitioning—are hierarchical and partitive approaches [19]. According to our trials both approaches gave the same results and we preferred to use partitive approach of clustering methods. Partitive clustering algorithms divide a data set into a number of clusters, typically by trying to minimize some criterion or error function. The number of clusters is usually predefined, but it can also be part of the error function [20]. The algorithm consists of the following steps given in Fig. 1 [19].

-
- Step 1. Determine the number of clusters
 - Step 2. Initialize the cluster centres
 - Step 3. Compute partitioning for data
 - Step 4. Compute (update) cluster centres
 - Step 5. If the partitioning is unchanged (or the algorithm has converged), stop. Otherwise, return to step 3
-

Figure 1. Steps of clustering algorithm

K-means clustering of implementations, simplicity, efficiency, and empirical success are the main reasons for its popularity [21]. The K-means algorithm was introduced MacQueen [22]. K-means is typically used with the Euclidean metric for computing the distance between points and cluster centres. As a result, K-means finds spherical or ball-shaped clusters in data. K-means with Mahalanobis distance metric has been used to detect hyper ellipsoidal clusters [23]. To put it simply, K-means finds k number of centroids, and then assigns all data points to the closest cluster, with the aim of keeping the centroids small.

There are a lot of applications of the K-means clustering, from unsupervised learning of neural network, pattern recognitions, classification analysis, artificial intelligence, image processing, etc. [21]. In principle, if there are several objects and each object have attributes and the aim is to classify the objects based on the attributes, then this algorithm can be applied [9].

3. EMPIRICAL RESULTS

This section includes the information about components of DEA and results of the efficiency analysis. In this study, 22 STPs are considered for the relevant analysis. These STPs are selected among the active STPs which are at least five years old. 14 of the STP are active for at least more than 10 years and eight of them are active at least five years. In addition to this, the data used in analysis belongs to year of 2017.

3.1. Variables and Data

STPs are evaluated using various DEA models to understand the weakness and strength of the DMUs. For this purpose, the variables used in DEA models are presented in Table 1. These variables are relevant to R&D activities of the firms in STPs. We used three input variables (number of personnel, number of firms, R&D expenditure) and three output variables (total R&D revenue, number of projects, number of patents).

Descriptive statistics of input and output variables are presented in Table 2. The table includes mean, standard deviation, minimum value and maximum value of the variables used in the analysis. The mean of the number of personnel in selected STPs varies between 111 and 9009. The number of firms in these STPs is between 37 and 395. The mean of R&D expenditures of firms in STPs is 173613880.70 TRY with a minimum value of 3707171 TRY and a maximum value of 838613194.10 TRY. The mean of revenue of firms in STPs is 307585614.60 TRY. Number of R&D projects in STPs varies between 41 and 1153. Minimum number of patents is 1 and maximum number of patents is 110 in STPs.

Also, DEA technique presumes the existence of a relationship among inputs and outputs data [24]. The correlation analysis of input and output variables is given in Table 3. According to pairwise correlation coefficients there is a positive and strong correlation between input and output variables used in DEA models. Coefficient values vary from 0.715 to 0.982

Table 1. Variables used in DEA

Variables	Abbreviation	Definition
<i>Input Variables</i>		
Number of personnel	PER	Total number of R&D and support personnel in STPs
Number of firms	FIRM	Total number of active R&D and incubation firms in STPs
R&D Expenditure	EXPD	Total R&D expenditures of firms in STPs
<i>Output Variables</i>		
Total R&D Revenue	REV	Total revenue obtained by the result of R&D projects developed by the firms in STPs
Number of Projects	PRJ	Total number of developed R&D projects by the firms in STPs
Number of Patents	PAT	Total number of grants and applications to national, international and triadic patents of firms in STPs

Table 2. Descriptive statistics of variables

	Mean	Std. Dev.	Min.	Max.
PER	2197.36	2785.61	111	9009
FIRM	161.96	101.07	37	395
EXPD	173613880.70	255124322.90	3707171	838613194.10
REV	307585614.60	489864987.10	3018134	2139462163
PRJ	374.77	306.27	41	1153
PAT	19.55	27.53	1	110

Table 3. Correlation coefficients between input and output variables

	PER	FIRM	EXPD	REV	PRJ	PAT
PER	1.000					
FIRM	0.916*	1.000				
EXPD	0.982*	0.855*	1.000			
REV	0.900*	0.824*	0.901*	1.000		
PRJ	0.953*	0.961*	0.912*	0.901*	1.000	
PAT	0.878*	0.752*	0.903*	0.715*	0.808*	1.000

3.2. Result of DEA

For the efficiency analysis of STPs, we used input-oriented DEA models. The input-oriented efficiency measures indicate the required improvement for the effective decision-making unit by comparing the actual input level with the best input level while keeping the output constants. We chose the input-orientation model since STPs have control on inputs rather than outputs. In Table 4, efficiency scores obtained by input-oriented CCR model is presented. The name of the STPs are coded since the data privacy restrictions.

According to the results in Table 4, there are five efficient and seventeen inefficient STPs. Average efficiency score of STPs is 0.840 and the lowest efficiency score is 0.562. In the third column of the table rank of the STPs is given.

3.3. Understanding Strong and Weak Areas of STPs

In this study, a DEA model with constant inputs is used with three inputs and three outputs as mentioned previous section. Including each input with all outputs and including each output with all inputs, we can obtain six more different DEA models and these models provide us to understand in which areas the STPs have good or bad performance. These results may be used for the policy-

making studies for STPs. These models are named as *Personnel Efficiency Model* (M_1), *Firm Efficiency Model* (M_2), *R&D Expenditure Efficiency Model* (M_3), *Revenue-oriented Model* (M_4), *Project-oriented Model* (M_5), and *Patent-oriented Model* (M_6).

In Table 5, the definition of these models is presented. In the first column of the table, the name of the model is given and in the following columns which variables are included in the model is specified.

In Table 6, efficiency scores of the corresponding DEA models are presented. According to the results, average efficiency score of *Constant Model* (M_0) is 0.840, *Personnel Efficiency Model* (M_1) is 0.680, *Firm Efficiency Model* (M_2) is 0.575, *R&D Expenditure Efficiency Model* (M_3) is 0.678, *Revenue-oriented Model* (M_4) is 0.451, *Project-oriented Model* (M_5) is 0.790 and *Patent-oriented Model* (M_6) is 0.438.

According to the results in Table 6, STP_14 is efficient in all models except *Patent-oriented Model* (M_6). But in this model, it has a high efficiency score (0.900). STP_11 is efficient in *Constant Model* (M_0) but it has low

Table 4. Efficiency scores of STPs

Name of STP	Efficiency Score	Rank	Name of STP	Efficiency Score	Rank
STP_11	1.000	1	STP_22	0.858	8
STP_12	1.000	1	STP_5	0.851	9
STP_14	1.000	1	STP_4	0.807	10
STP_15	1.000	1	STP_17	0.734	11
STP_19	1.000	1	STP_21	0.719	12
STP_9	0.987	2	STP_3	0.703	13
STP_16	0.981	3	STP_2	0.679	14
STP_18	0.973	4	STP_6	0.675	15
STP_1	0.962	5	STP_20	0.585	16
STP_7	0.931	6	STP_10	0.570	17
STP_13	0.905	7	STP_8	0.562	18
Average	0.840				

Table 5. DEA Models for the STPs

	Inputs			Outputs		
	PER	FIRM	EXPD	REV	PRJ	PAT
Constant Model (M_0)	•	•	•	•	•	•
Personnel Efficiency Model (M_1)	•			•	•	•
Firm Efficiency Model (M_2)		•		•	•	•
R&D Expenditure Efficiency Model (M_3)			•	•	•	•
Revenue-oriented Model (M_4)	•	•	•	•		
Project-oriented Model (M_5)	•	•	•		•	
Patent-oriented Model (M_6)	•	•	•			•

Table 6. Efficiency scores of STPs in terms of DEA models

Name of STP	M_0	M_1	M_2	M_3	M_4	M_5	M_6
STP_1	0.962	0.540	0.347	0.952	0.689	0.925	0.578
STP_2	0.679	0.623	0.429	0.679	0.541	0.679	0.201
STP_3	0.703	0.702	0.534	0.600	0.295	0.703	0.455
STP_4	0.807	0.718	0.192	0.807	0.299	0.807	0.115
STP_5	0.851	0.780	0.669	0.335	0.137	0.851	0.430
STP_6	0.675	0.648	0.650	0.544	0.470	0.675	0.491
STP_7	0.931	0.895	0.522	0.704	0.209	0.931	0.252
STP_8	0.562	0.425	0.399	0.558	0.305	0.562	0.130
STP_9	0.987	0.795	0.514	0.987	0.881	0.981	0.213
STP_10	0.570	0.339	0.267	0.570	0.428	0.564	0.064
STP_11	1.000	1.000	1.000	0.484	0.318	0.484	1.000
STP_12	1.000	0.510	0.270	0.917	0.575	0.889	1.000
STP_13	0.905	0.905	0.823	0.619	0.546	0.645	0.857
STP_14	1.000	1.000	1.000	1.000	1.000	1.000	0.900
STP_15	1.000	0.761	0.509	1.000	1.000	0.923	0.406
STP_16	0.981	0.787	0.981	0.455	0.168	0.981	0.204
STP_17	0.734	0.542	0.734	0.433	0.342	0.671	0.444
STP_18	0.973	0.875	0.771	0.563	0.224	0.973	0.446
STP_19	1.000	0.852	1.000	0.613	0.177	1.000	0.483
STP_20	0.585	0.442	0.277	0.585	0.529	0.572	0.305
STP_21	0.719	0.505	0.485	0.654	0.310	0.719	0.114
STP_22	0.858	0.310	0.267	0.858	0.471	0.847	0.549
Average	0.840	0.680	0.575	0.678	0.451	0.790	0.438

efficiency performance in *R&D Expenditure Efficiency Model* (M_3) *Revenue-oriented Model* (M_4) and *Project-oriented Model* (M_5). This means STP_11 is not good at transforming the R&D expenditures to projects and revenues. STP_15 is efficient in *Constant Model* (M_0) but it has a low performance in *R&D Expenditure Efficiency Model* (M_3) and *Patent-oriented Model* (M_6). This means, it produces less output by R&D expenditures and outputs of the projects are not applied or granted as a

patent in enough level. Another efficient unit STP_19 have a low performance in *Revenue-oriented Model* (M_4) and *Patent-oriented Model* (M_6). This can be interpreted as outputs of the R&D projects in STP_19 do not commercialize and do not get patents in enough level.

Personnel Efficiency Model (M_1), *Firm Efficiency Model* (M_2) and *R&D Expenditure Efficiency Model* (M_3) scores are close to each other and these values are not far from the *Constant Model* (M_0). In other words, input efficiency

scores do not fluctuate more. Efficiency score of *Project-oriented Efficiency Model* (M_5) is 0.790 and this means STPs generate R&D projects using their inputs effectively. But *Revenue-oriented Efficiency Model* (M_4) and *Patent-oriented Efficiency Model* (M_6) scores decrease more than other scores. This result can be interpreted as, there is a problem in the commercialization and in the transforming to patents of the project results. To increase the performance in this area, firms in STPs may have mentoring on commercialization and intellectual property rights. Furthermore, the competitiveness of the R&D projects also should be examined. If the R&D projects are not innovative enough their chance to commercialized and to be subject to patents will decrease.

To understand if there is a significant difference between the DEA models Wilcoxon signed-rank test is conducted and the results of this test is given in Table 7. According to Wilcoxon signed-rank test results for DEA models, difference between the constant model and other models are significant at the 1% confidence level as presented in the last column of the table.

3.4. Findings of Cluster Analysis

In the study, we clustered the STPs based on the efficiency scores of the models obtained in the previous section. The variables included in the clustering is selected according to the pairwise correlation results given in Table 8. According to the table, there is a significant positive correlation between M_1 - M_2 , M_3 - M_4 and M_3 - M_5 models in the significance levels of 1% and 5%.. The variables having no correlation with each other are used in K-means clustering analysis. For this reason, efficiency scores of M_1 , M_4 , M_5 and M_6 models are included in K-means clustering analysis. M_1 is an input-related model and M_4 , M_5 and M_6 are output-related models.

ANOVA table of K-means algorithm is given in Table 9. According to the table, clusters are different in terms of the variables and three clusters are obtained by the algorithm.

There are four STPs in Cluster 1, five STPs in Cluster 2 and 13 STPs in Cluster 3. The name of STPs in each cluster is given in Table 10.

In Table 11 mean values of efficiency models by clusters are given. For M_1 Cluster 3 performs best, for M_4 Cluster 1 performs best, for M_5 Cluster 1 performs best and for M_6 Cluster 3 performs best. Cluster 1 STPs are named as *Marketers*, Cluster 2 STPs are named as *Researchers* and Cluster 3 STPs are named as *Low-performers* by the authors.

Marketers are best at the gathering revenue and developing R&D projects. They have moderate performance at personnel efficiency and gathering patents. STPs in this group develops effective R&D projects and they commercialize these projects. In intellectual property rights are they need to be improved.

Researchers are best at personnel efficiency and patents. They have moderate performance at revenues and R&D projects. STPs in this group have good quality of R&D personnel and researchers, so they can gather patents. But in commercialization area they need to be improved.

Low-performers have the lowest performance in all areas. The main source of this situation is these STPs are in less-developed regions so the infrastructure in research and industrialization is not good enough. To improve the performance of these STPs investments in universities and industry must be increased in these regions and collaborations with the developed STPs must be encouraged in *Marketers* and *Researchers* clusters.

4. CONCLUSIONS

In this study, the efficiency of selected STPs in Turkey are examined using DEA. The DEA model includes three inputs and three outputs for 22 STPs. For the analysis STPs that are at least five years old ones selected. In the first stage of the study general DEA model is constructed using all input and output variables. In the second stage to understand the strength and weakness of the STPs six additional DEA models are constructed. In the last stage these STPs are clustered based on efficiency scores of the DEA models.

According to the results, there five efficient STPs and average efficiency score of the STPs is 0.840. In further analysis, we see that STPs have better performance in input-related models than output-related models. Especially STPs have low performance in R&D revenue-oriented efficiency model and patent-oriented model. This result can be interpreted as, there is a problem in the commercialization and in the transforming to patents of the projects. Furthermore, the competitiveness of the R&D projects also should be examined. If the R&D projects are not innovative enough their chance to commercializing and to be subject to patents will decrease. In clustering stage STPs are clustered using K-means algorithm using *Personnel Efficiency Model* (M_1), *Revenue-oriented Model* (M_4), *Project-oriented Model* (M_5) and *Patent-oriented Model* (M_6). At the end of clustering three clusters are obtained and these clusters are named as *Marketers*, *Researchers* and *Low-performers*. The features of the STPs in each cluster was interpreted and suggestions were made to improve the performance of the STPs in each cluster.

For the future directions, additional input and output variables can be used such as R&D personnel, R&D investments, exports etc. Add some examples and other efficiency analysing methods such as network DEA models or stochastic frontier analysis can be used. Furthermore, regression models can be used to examine the determinants on the efficiency of the STPs as the second stage analysis.

DECLARATION OF ETHICAL STANDARDS

The author(s) of this article declare that the materials and methods used in this study do not require ethical committee permission and/or legal-special permission.

AUTHORS' CONTRIBUTIONS

Necla ARSLAN: Established the models and analyzed the results.

Table 7. Wilcoxon signed-rank test results for DEA models output variables

		N	Mean Rank	Sum of Ranks	z	P
M ₀ -M ₁	Negative	0	0,00	,00	-3,823	0,000*
	Positive	19	10,00	190,00		
	Ties	3				
M ₀ -M ₂	Negative	0	0,00	,00	-3,622	0,000*
	Positive	17	9,00	153,00		
	Ties	5				
M ₀ -M ₃	Negative	0	0,00	,00	-3,297	0,001*
	Positive	14	7,50	105,00		
	Ties	8				
M ₀ -M ₄	Negative	0	0,00	,00	-3,920	0,000*
	Positive	20	10,50	210,00		
	Ties	2				
M ₀ -M ₅	Negative	0	0,00	,00	-2,805	0,005*
	Positive	10	5,50	55,00		
	Ties	12				
M ₀ -M ₆	Negative	0	0,00	,00	-3,920	0,000*
	Positive	20	10,50	210,00		
	Ties	2				

*p<0.01

Table 8. Correlation coefficients between DEA models

	M ₀	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆
M ₀	1.000						
M ₁	0,663**	1.000					
M ₂	0,484*	0,757**	1.000				
M ₃	0,413	-0,009	-0,339	1.000			
M ₄	0,729**	0,369	0,206	0,489*	1.000		
M ₅	0,233	0,030	-0,137	0,795**	0,201	1.000	
M ₆	0,542**	0,384	0,408	0,184	0,057	0,277	1.000

Table 9. ANOVA results of K-means algorithm

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
M ₁	0.121	2	0.034	19	3.562	0.049**
M ₄	0.075	2	0.023	19	3.298	0.059***
M ₅	0.478	2	0.022	19	22.151	0.000*
M ₆	0.414	2	0.047	19	8.865	0.002*

*p<0.01, **p<0.05, ***p<0.1

Table 10. Efficiency score clusters

Cluster	Name of STPs
1	STP_1, STP_14, STP_23, STP_24
2	STP_17, STP_18, STP_19, STP_27, STP_28
3	STP_2, STP_3, STP_4, STP_6, STP_8, STP_10, STP_12, STP_16, STP_25, STP_26, STP_30, STP_31, STP_32

Table 11. Mean values of efficiency models by clusters

Efficiency Model	Cluster 1 (Marketers)	Cluster 2 (Researchers)	Cluster 3 (Low-performers)
M ₁	0.774	0.828	0.594
M ₄	0.957	0.798	0.735
M ₅	0.893	0.368	0.347
M ₆	0.524	0.757	0.289

Önder BELGİN: Constructed the conceptual framework and wrote the manuscript.

CONFLICT OF INTEREST

There is no conflict of interest in this study.

REFERENCES

- [1] Hobbs, K.G., Link, A.N. and Scott, J.T., “Science and technology parks: an annotated and analytical literature review”, *Journal of Technology Transfer*, 42: 957–76, (2017).
- [2] Hu, J-L., Yeh, F-Y. and Chang, I-T., “Industrial park efficiency in Taiwan”, *Journal of Information and Optimization Sciences*, 30(1): 63-86, (2009).
- [3] Hu, J-L., Han, T-F., Yeh, F-Y. and Lu, C-L., “Efficiency of science and technology industrial parks in China”, *Journal of Management Research*, 10(3): 151-166, (2010).
- [4] Farrell, M. J., “The measurement of productive efficiency”, *Journal of the Royal Statistical Society*, 120: 253-290, (1957).
- [5] Charnes, A., Cooper, W. W. and Rhodes, E., “Measuring the efficiency of decision-making units”, *European Journal of Operational Research*, 2: 429-444, (1978).
- [6] Banker R.D., Charnes, A. and Cooper, W.W., “Some models for estimating technical and scale inefficiencies in data envelopment analysis”, *Management Science*, 30(9): 1078-1092, (1984).
- [7] Hirschberg, J.G. and Lye, J.N., “Clustering in data envelopment analysis using bootstrapped efficiency scores”, *Papers 800*, Department of Economics, Melbourne, (2001).
- [8] Sherman, H.D. and Zhu, J., “Service productivity management: Improving service performance using data Envelopment analysis (DEA)”, *Springer*, New York, (2006).
- [9] Lemos, C.A.A., Lins, M.P.E. and Ebecken, N.F.F., “DEA implementation and clustering analysis using the K-means algorithm”, Data Mining VI, *WIT Press*, Brazil, (2005).
- [10] Bougnol M-L., Dula, J.H., EstellitaLins, M.P. and Moreira da Silva, A.C., “Enhancing standard performance practices with DEA”, *Omega*, 38(1–2): 33-45, (2010).
- [11] Maghyreh, A. I. and Awartani, B., “The effect of market structure, regulation, and risk on banks efficiency: Evidence from the Gulf cooperation council countries”, *Journal of Economic Studies*, 41(3): 405-430, (2014).
- [12] Cook, W.D., Seiford, L.M., “Data envelopment analysis (DEA) – thirty years on”, *European Journal of Operations Research*, 192(1): 1–17, (2009).
- [13] Liu, J.S., Lu, L.Y.Y., Lu, W.-M., Lin, B.J.Y., “Data envelopment analysis 1978–2010: a citation-based literature survey”. *Omega*, 41(1): 3–15, (2013).
- [14] Liu, J.S., Lu, L.Y.Y., Lu, W.-M., Lin, B.J.Y., “A survey of DEA applications”, *Omega*, 41(5): 893–902, (2013).
- [15] Fraley, C. and Raftery, A.E., “Model-based clustering, discriminant analysis, and density estimation”, *Journal of the American Statistical Association*, 97(458): 611-631, (2002).
- [16] Alpaydin, E., “Introduction to machine learning”, *The MIT Press*, London, (2014).
- [17] Jain, A., Murty, M. and Flynn, P. “Data clustering: A review”, *ACM Computing Surveys*, 31: 264-323, (1999).
- [18] Han B, Liu L. and Omiecinski E., “NEAT: Road network aware trajectory clustering”, *Proceedings of the 32nd IEEE International Conference on Distributed Computing Systems*, Macau, China, 142–151, (2012).
- [19] Vesanto, J., and Alhoniemi, E., “Clustering of the self-organizing map”, *IEEE Transactions on Neural Networks*, 11(3): 586-600, (2000).
- [20] Buhmann, J. and Kühnel, H., “Complexity optimized data clustering by competitive neural network”, *Neural Computation*, 5(1): 75-88, (1993).
- [21] Jain, A.K., “Data clustering: 50 years beyond K-means”, *Pattern Recognition Letters*, 31: 651-666, (2010).
- [22] MacQueen, J.B. ‘Some methods for classification and analysis of multivariate observations’, *Proceedings of 5th Berkley Symposium on Mathematical Statistics and Probability*, I, California, 281-297, (1967).
- [23] Mao, J. and Jain, A.K., “A self-organizing network for hyper-ellipsoidal clustering (HEC)”, *Transactions on Neural Networks*, 7(1): 16-29, (1996).
- [24] Hung, S.-W., Wang, A.P., “Entrepreneurs with Glamour? DEA Performance Characterization of High-Tech and Older-Established Industries”, *Economic Modelling*, 29: 1146–1153, (2012).