

RESEARCH ARTICLE

## Copula approach to select input/output variables for DEA

Olcay Alpay<sup>a\*</sup>, Elvan Akturk Hayat<sup>b</sup>

<sup>a</sup> Department of Statistics, Sinop University, Turkey

<sup>b</sup> Department of Econometrics, Adnan Menderes University, Turkey  
[olcayb@sinop.edu.tr](mailto:olcayb@sinop.edu.tr), [elvan.hayat@adu.edu.tr](mailto:elvan.hayat@adu.edu.tr)

---

### ARTICLE INFO

#### Article history:

Received: 5 April 2016

Accepted: 15 August 2016

Available Online: 25 October 2016

#### Keywords:

Data envelopment analysis

Variable selection

Copulas

Local dependence function

AMS Classification 2010:

90C30, 60E99

### ABSTRACT

Determination of the input/output variables is an important issue in Data Envelopment Analysis (DEA). Researchers often refer to expert opinions in defining these variables. The purpose of this paper is to propose a new approach to determine the input/output variables, it is important to keep in mind that especially when there is no any priori information about variable selection. This new proposed technique is based on a theoretical method which is called "Copula". Copula functions are used for modeling the dependency structure of the variables with each other. Also we use the local dependence function which analyzes the point dependency of variables of copulas to define the input/output variables. To illustrate the usefulness of the proposed approach, we conduct two applications using simulated and real data and compare the efficiencies in DEA. Our results show that new approach gives values close to perfection.



## 1. Introduction

Data Envelopment Analysis (DEA) is a data oriented non-parametric method, introduced by Charnes et al. [1] to evaluate the relative efficiency of organizational units called as decision making units (DMUs) according to selected input and output variables [3]. Manufacturing units, departments of schools or hospitals, a set of firms or even practising individuals can be given as examples to DMUs [7].

Nowadays, institutions producing service or product, are obliged to have an effective performance because of intense competitive conditions and limited sources. Although efficiency of the DMU's depends crucially on input/output variables to be selected correctly, there is no any guidelines or systematic procedure to classify the variables. Many researchers often make such determination subjectively or using expert opinions. Some of them use the variables based directly on other's selection results according to calculated efficiency [8].

The issue of variable selection seems to be a subject rarely studied in the literature. Considering the studies over the last 15 years, it can be seen that Pastor et al. [9] defined a measure, namely ECM (efficiency contribution measure), to evaluate the input or output variables. In 2003, Jenkins and Anderson [10]

proposed a systematic statistical approach to reduce the number of already defined input and output variables. Ruggiero [11] addressed the subject of input selection using regression analysis. Edirisinghe and Zhang [12] proposed a Generalized DEA approach to determine the type of variable via maximizing the correlation between DEA-based score of financial strength and stock market performance. In 2009, Morita and Avkiran [3] proposed a selection method utilizing diagonal layout design. Finally, Madhanagopal and Chandrasekaran [13] developed a genetic algorithm approach for the selection of input/output variables.

The approach we propose in this study takes into consideration distribution of data and the point dependency between candidate variables, which makes it distinctive amongst the others. Copulas and the local dependence function were used to achieve the correct discrimination of input/output variables. We assume that there is no any information or expert opinion to decide the input/output variables. Performance assessment of our proposed method was made by means of simulation, followed by a real data application. In this way, we were able to make comparison of the efficiencies of the new approach and those of the one based directly on expert opinion selection method for this study. The main goal of this

---

\*Corresponding author

paper is to propose a new theory based approach to determine the input/output variables.

The paper is organized as follows: Section 2 gives a brief summary of the DEA models. In section 3 we define copulas (especially Farlie-Gumbel-Morgenstern “FGM” copula) and local dependence function. In this section an algorithm is also developed to determine the input/output variables. In Section 4, a simulation and a real data applications are conducted and results are evaluated. The last section gives the conclusions.

## 2. Data envelopment analysis (DEA)

Data envelopment analysis (DEA) is a linear programming based technique to calculate the relative efficiencies of a set of decision making units (DMUs) that have similar inputs and similar outputs. Also it is a non-parametric method as it does not require any assumption about functional form. This technique aims to measure how efficiently a DMU uses the resources available to generate a set of outputs [1]. DMUs can include manufacturing units, departments of big organizations such as universities, schools, bank branches, hospitals, power plants, police stations, tax offices, prisons etc.

The basic frontier model was developed by Charnes et al. [1] known as the CCR model, but now widely called as CRS (constant returns to scale) and was extended by Banker et al. [14] to include variable returns to scale (VRS). So the basic DEA models are known as CCR and BCC referred to as the VRS.

DEA models have two orientations: input-oriented and output-oriented. With input-oriented DEA, the linear programming model is constituted so as to determine how much the input use of a firm could contract if used efficiently in order to achieve the same output level. In contrast, with output-oriented DEA, the linear programme is constituted to determine a firm’s potential output given its inputs if it operated efficiently as firms along the best practice frontier [15]. The input oriented model contracts the inputs as far as possible while controlling the outputs. The output oriented model expands the outputs as far as possible while controlling the inputs [16].

The original fractional CRS model Eq.(1) evaluates the relative efficiencies of  $n$  DMUs  $j=1, \dots, n$  each with  $m$  inputs and  $s$  outputs denoted by  $x_{1j}, x_{2j}, \dots, x_{mj}$  and  $y_{1j}, y_{2j}, \dots, y_{mj}$  respectively [1]. This is done so by maximizing the ratio of weighted sum of output to the weighted sum of inputs:

$$Max\theta_o = \frac{\sum_{r=1}^s u_r y_{ro}}{\sum_{i=1}^m v_i x_{io}}$$

s.t.

$$\frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1, j = 1, \dots, n, r = 1, \dots, s \quad (1)$$

$$u_r, v_i \geq 0 \text{ for all } r \text{ and } i$$

In model Eq. (1),  $\theta_o$  is the efficiency of  $DMU_o$  and  $u_r$  and  $v_i$  are the factor weights. For computational convenience the fractional form Eq.(1) is re-expressed in linear program (LP) form as follows which is known as input oriented CRS model:

$$Max\theta_o = \sum_{r=1}^s u_r y_{ro}$$

s.t.

$$\sum_{r=1}^s u_r y_{rj} - \sum_{i=1}^m v_i x_{ij} \leq 0, j = 1, \dots, n \quad (2)$$

$$\sum_{i=1}^m v_i x_{io} = 1$$

$$u_r, v_i \geq 0$$

The dual of linear program (LP) model for input oriented CRS model is as follows:

$$Min\theta_o$$

s.t.

$$\theta x_{io} \geq \sum_{j=1}^n \lambda_j x_{ij} \quad i = 1, \dots, m \quad (3)$$

$$y_{ro} \leq \sum_{j=1}^n \lambda_j y_{rj} \quad r = 1, \dots, s$$

$$\lambda_j \geq 0$$

The output-oriented CRS model is as follows:

$$Max\theta_o$$

s.t.

$$x_{io} \geq \sum_{j=1}^n \lambda_j x_{ij} \quad i = 1, \dots, m \quad (4)$$

$$\theta y_{ro} \leq \sum_{j=1}^n \lambda_j y_{rj} \quad r = 1, \dots, s$$

$$\lambda_j \geq 0$$

The CRS models (dual and primal) with input orientation are still the most widely known and used DEA models despite the numerous modified models that have appeared [17].

## 3. Copula, local dependence function and selection procedure

In general, copula is a function which helps re-define the joint distribution function using marginal distribution functions in  $I^2$  when the random variables are dependent. In recent years, copulas have been involved in many studies such as statistics, economics, finance and risk management, dependence measuring, modeling, and serial dependence in time series [18].

**Definition 1.** Let  $C(u,v)$  is defined as a bivariate function in  $I^2=[0,1] \times [0,1]$ . If this function has the following properties, it is called as two-dimensional copula function.

- $C(u,0)=C(0,v)=0$
- $C(u,1)=u$  and  $C(1,v)=v, \forall (u,v) \in I$

For every  $0 \leq u_1 \leq u_2 \leq 1$  and  $0 \leq v_1 \leq v_2 \leq 1$

- $V_c([u,v])=C(u_1,v_1)-C(u_1,v_2)-C(u_2,v_1)+C(u_2,v_2) \geq 0$

[4].

### 3.1. Local dependence function (LDF)

Let  $X$  and  $Y$  be random variables with marginal distribution functions  $F(x)$ ,  $F(y)$  and marginal probability density functions  $f(x)$  and  $f(y)$  respectively. The following function is obtained from the expression of the Pearson correlation coefficient by replacing mathematical expected values  $E(X)$  and  $E(Y)$  conditional expected values  $E(X|Y=y)$  and  $E(Y|X=x)$  [6].

$$H(x, y) = \frac{E\{(X - E(X|Y=y))(Y - E(Y|X=x))\}}{\sqrt{E\{(X - E(X|Y=y))^2\}}\sqrt{E\{(Y - E(Y|X=x))^2\}}} \tag{5}$$

$H(x,y)$  can be referred as a local dependence function which characterizes the dependence between  $X$  and  $Y$  at the point  $(x,y)$ . After simple mathematical transformation  $\xi_x = EX - E(X|Y=y)$  and  $\xi_y = EY - E(Y|X=x)$  the equation Eq. (5) can be written as

$$H(x, y) = \frac{\rho + \Phi_x(y)\Phi_y(x)}{\sqrt{1 + \Phi_x^2(y)}\sqrt{1 + \Phi_y^2(x)}}$$

where

$$\rho = \frac{Cov(X, Y)}{\sigma_x \sigma_y}$$

$$\Phi_x(y) = \frac{\xi_x(y)}{\sigma_x} \quad \Phi_y(x) = \frac{\xi_y(x)}{\sigma_y}$$

[5].

Local dependence function has the following properties

1. If  $X$  and  $Y$  are independent, then  $H(x,y)=0$  for all  $(x,y) \in N_{X,Y}$ .
2.  $|H(x,y)| \leq 1$ , for all  $(x,y) \in N_{X,Y}$ .
3. If  $|H(x,y)|=1$  for some  $(x,y) \in N_{X,Y}$ , then  $\rho \neq 0$ .
4. If  $\rho = \pm 1$ , then  $H(x,y) = \pm 1$ .
5. If  $H(x,y)=0$  for all  $(x,y) \in N_{X,Y}$ , then either  $E(X|Y=y)$  or  $E(Y|X=x)$  for all  $(x,y) \in N_{X,Y}$  and  $\rho=0$ .

[6].

### 3.2. FGM copula

Farlie-Gumbel-Morgenstern (FGM) distribution was introduced by Morgenstern in 1956 with Cauchy marginal. This class was examined by Gumbel for exponential marginal and further generalized by Farlie in 1960 [19].

Let  $(X,Y)$  be absolutely continuous random variables. The general distribution function is defined as

$$F(x, y) = F(x)G(y)\{1 + \alpha A(F(x))B(G(y))\} \tag{6}$$

where  $A(x) \rightarrow 0$ ,  $B(y) \rightarrow 0$  as  $x \rightarrow 1$ ,  $A(x)$  and  $B(y)$  satisfy certain regularity conditions ensuring that the Eq. (6) is a distribution function with absolutely continuous marginal  $F(x)$  and  $G(y)$  [20].

FGM copula is a positive quadrant dependent (PQD) copula and the local dependence function is as

follows;

$$H(x, y) = \frac{\alpha/3 + (\alpha^2/3)(1-2x)(1-2y)}{\sqrt{[1 + (\alpha^2/3)(1-2x)^2][1 + (\alpha^2/3)(1-2y)^2]}}$$

where  $\alpha$  is the association parameter [5].

### 3.3. Quadrant dependence

**Definition 2.** Let  $X$  and  $Y$  be random variables.  $X$  and  $Y$  are positive quadrant dependent (PQD) if all  $(x,y)$  variables in  $R^2$

$$P\{X \leq x, Y \leq y\} \geq P\{X \leq x\}P\{Y \leq y\}$$

or equivalently

$$P\{X > x, Y > y\} \geq P\{X > x\}P\{Y > y\}$$

Similarly,  $X$  and  $Y$  are negative quadrant dependent (NQD) if

$$P\{X \leq x, Y \leq y\} < P\{X \leq x\}P\{Y \leq y\}$$

or equivalently

$$P\{X > x, Y > y\} < P\{X > x\}P\{Y > y\}$$

[4].

### 3.4. Selection procedure

In this part of the study, we present a new algorithm to determine the input/output variables. The following algorithm explained in detail leads to the selection process.

**Step-1:** Determine the distribution of variables and construct the appropriate copula function (e.g. if the distribution is uniform the FGM is suitable).

**Step-2:** Construct the LDF of determined copula according to the method as described in Section 3.1.

**Step-3:** Determine the type of quadrant dependency of the copula mentioned in Section 3.3., to use to decide which group is selected as input/output.

**Step-4:** Calculate the LDF values for pairwise variables. The variables which have  $Max|H(X_i, X_j)|$ , selected as reference variables ( $x_{r1}$  and  $x_{r2}$ ) and assign them two separate groups (*Group-1* and *Group-2*). Allocate the rest of variables to the groups according to the following procedure:

If  $|H(x_i, x_{r1})| > |H(x_i, x_{r2})|$  (where  $x_i \neq x_{r1}$  and  $x_i \neq x_{r2}$ ) then  $x_i$  is assigned to the Group-1 else it is assigned to the Group-2.

**Step-5:** Make predetermination about whether the study is input or output oriented.

**Step-6:** Determine the sign of LDFs which is calculated between the each of the rest of the variables and the reference variables for two separate groups constructed in Step-4.

*a. Suppose that study is input oriented and copula is PQD,*

*i. If one of the groups has only positive LDFs and other has only negative, then choose the variables in group which has positive LDFs as input variables.*

*ii. If all of LDFs are positive in each group, then select the variables in group which has  $MaxH(X_i, X_j)$*

as input variables. If all of LDFs are negative in each group, then select the variables in group which has  $Min/H(X_i, X_j)/$  as input variables.

*iii.* If groups contain different sign LDFs, then select the variables in group which has  $Min/H(X_i, X_j)/$  as input variables.

*b. Suppose that study is input oriented and copula is NQD,*

*i.* If one of the groups has only positive LDFs and other has only negative, then choose the variables in group which has negative LDFs as input variables.

*ii.* If all of LDFs are positive in each group, then select the variables in group which has  $MinH(X_i, X_j)$  as input variables. If all of LDFs are negative in each group, then select the variables in group which has  $MaxH(X_i, X_j)/$  as input variables.

*iii.* If groups contain different sign LDFs, then select the variables in group which has  $Min/H(X_i, X_j)/$  as input variables.

*c. Suppose that study is output oriented and copula is PQD,*

*i.* If one of the groups has only positive LDFs and other has only negative, then choose the variables in group which has positive LDFs as output variables.

*ii.* If all of LDFs are positive in each group, then select the variables in group which has  $MaxH(X_i, X_j)$  as output variables. If all of LDFs are negative in each group, then select the variables in group which has  $Min/H(X_i, X_j)/$  as output variables.

*iii.* If groups contain different sign LDFs, then select the variables in group which has  $Min/H(X_i, X_j)/$  as output variables.

*d. Suppose that study is output oriented and copula is NQD,*

*i.* If one of the groups has only positive LDFs and other has only negative, then choose the variables in group which has negative LDFs as output variables.

*ii.* If all of LDFs are positive in each group, then select the variables in group which has  $MinH(X_i, X_j)$  as output variables. If all of LDFs are negative in each group, then select the variables in group which has  $Max/H(X_i, X_j)/$  as output variables.

*iii.* If groups contain different sign LDFs, then select the variables in group which has  $Min/H(X_i, X_j)/$  as output variables.

**Step-7:** Run the DEA model and calculate the efficiencies.

**4. Application**

In this part of the study, a simulation and a real data example have been conducted and DEA have been performed using the package program DEAP 2.1. The data for simulation study were generated from Uniform (0,1) and FGM copula was selected to represent them. We used the Yoluk’s [2] data set for the real data example.

**4.1. Simulation study**

In simulation part of the study, we use five variables and 20 decision making units. Table 1 shows the min, max and absolute values of LDFs.

**Table 1.** Values of local dependence function.

Pairs of variables	$MinH(X_i, X_j)$	$MaxH(X_i, X_j)$	$ H(X_i, X_j) $
$X_1-X_2$	0.0613	0.0845	0.0845
$X_1-X_3$	0.1111	0.2274	0.2274
$X_1-X_4$	0.0367	0.0431	0.0431
$X_1-X_5$	-0.0083	-0.0080	0.0083
$X_2-X_3$	0.0668	0.0956	0.0956
$X_2-X_4$	0.1348	0.2156	0.2156
$X_2-X_5$	-0.0851	-0.3029	0.3029
$X_3-X_4$	0.1482	0.3355	0.3355
$X_3-X_5$	0.2226	0.0980	0.2226
$X_4-X_5$	-0.4272	-0.1590	0.4272

Maximum local dependency value is between  $X_4$  and  $X_5$  ( $Max/H(X_i, X_j)=0.4272$ ). So these variables are selected as reference. Allocated variables according to descending of  $Max/H(X_i, X_j)/$  are given below.

Group 1	Group 2
$X_4$	$X_5$
$X_3$	$X_2$
$X_1$	

$H(X_2, X_5) < 0$  and  $H(X_1, X_4), H(X_3, X_5) > 0$

Suppose that the study is input oriented. So variables in Group 1 are the input variables. DEA results as follows

**Table 2.** Efficiency scores.

DMU	Efficiency	DMU	Efficiency
1	0.144	11	0.428
2	0.213	12	0.472
3	1.000	13	1.000
4	1.000	14	1.000
5	0.905	15	0.755
6	0.322	16	0.367
7	0.688	17	0.136
8	0.521	18	1.000
9	0.306	19	1.000
10	0.840	20	0.883

**4.2. Real data application**

In this part of the study we use Yoluk’s [2] hospital data given in Table 3. The efficiency analysis was performed as input oriented and the first three variables were taken as input, the last four variables were taken as output variables. We assume that we have no any information about the variables and are not able to get an expert opinion. Firstly, we tested the goodness of fit for all variables to the Uniform distribution and found no departures from the hypothesized distribution. At the second stage the LDF values for FGM copula were computed on these variables standardized to Uniform (0,1) and results

were given in Table 4. Variables have been classified with these values according to the proposed algorithm. The efficiency scores of our approach and Yoluk's results are given in Table 5.

**Table 3.** Values of real data.

DMU	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>
1	779	360	599	1691527	33685	52095	199291
2	776	421	807	1510740	40175	52890	266637
3	1090	544	1092	1231625	61132	68642	376193
4	680	431	602	1545480	32376	26424	171229
5	184	193	308	765675	15236	14119	78468
6	142	231	401	414665	11360	10018	65483
7	1105	989	678	555320	28384	42653	214864
8	968	425	1003	779960	63300	36235	314584
9	1039	951	2111	736499	53657	66304	603414

**Table 4.** Values of local dependence function for real data.

Pairs of variables	MinH(X <sub>i</sub> ,X <sub>j</sub> )	MaxH(X <sub>i</sub> ,X <sub>j</sub> )	H(X <sub>i</sub> ,X <sub>j</sub> )
X <sub>1</sub> -X <sub>2</sub>	0.9156	0.1473	0.9156
X <sub>1</sub> -X <sub>3</sub>	0.8312	-0.0417	0.8312
X <sub>1</sub> -X <sub>4</sub>	0.2927	0.0969	0.2927
X <sub>1</sub> -X <sub>5</sub>	0.9337	0.0759	0.9337
X <sub>1</sub> -X <sub>6</sub>	0.9482	0.4697	0.9482
X <sub>1</sub> -X <sub>7</sub>	-0.8897	-0.9853	0.9853
X <sub>2</sub> -X <sub>3</sub>	0.8634	-0.0641	0.8634
X <sub>2</sub> -X <sub>4</sub>	-0.0743	-0.3236	0.3236
X <sub>2</sub> -X <sub>5</sub>	0.5622	0.1586	0.5622
X <sub>2</sub> -X <sub>6</sub>	0.7964	0.2526	0.7964
X <sub>2</sub> -X <sub>7</sub>	-0.1652	-0.8225	0.8225
X <sub>3</sub> -X <sub>4</sub>	-0.0660	-0.1027	0.1027
X <sub>3</sub> -X <sub>5</sub>	0.8775	0.2293	0.8775
X <sub>3</sub> -X <sub>6</sub>	0.8874	0.1683	0.8874
X <sub>3</sub> -X <sub>7</sub>	-0.1369	-0.8711	0.8711
X <sub>4</sub> -X <sub>5</sub>	0.3282	0.1484	0.3282
X <sub>4</sub> -X <sub>6</sub>	0.5607	0.1477	0.5607
X <sub>4</sub> -X <sub>7</sub>	-0.1207	-0.6102	0.6102
X <sub>5</sub> -X <sub>6</sub>	0.9109	0.3518	0.9109
X <sub>5</sub> -X <sub>7</sub>	-0.2999	-0.9562	0.9562
X <sub>6</sub> -X <sub>7</sub>	-0.3542	-0.9614	0.9614

X<sub>1</sub> and X<sub>7</sub> are reference variables and the rest of variables are allocated as given below.

Group 1	Group 2
X <sub>1</sub>	X <sub>7</sub>
X <sub>2</sub>	X <sub>6</sub>
	X <sub>5</sub>
	X <sub>4</sub>
	X <sub>3</sub>

$$H(X_1, X_2) > 0 \text{ and } H(X_3, X_7), H(X_4, X_7), H(X_5, X_7), \\ H(X_6, X_7) < 0$$

Variables in Group 1 should be selected as input variables. As seen from the table, the copula approach has almost done exact classifying and has determined input/output variables correctly.

**Table 5.** The efficiency scores of real data.

DMU	Efficiency <sup>§</sup>	Efficiency*
1	1.00	1.00
2	1.00	1.00
3	1.00	1.00
4	1.00	1.00
5	1.00	1.00
6	1.00	1.00
7	0.50	0.78
8	1.00	1.00
9	1.00	1.00

<sup>§</sup> Copula approach's efficiency scores and

\* Yoluk's efficiency scores.

### 5. Conclusion

Data Envelopment Analysis is the most frequently used method to evaluate the efficiencies of DMUs. Although the determination of input/output variables is one of the most important problem of DEA, there have been limited attempts in the literature to solve this issue. Getting an expert opinion, which is the most preferred method, may actually be a subjective approach and it may also become an expensive way to identify the variables as input/output. The methods to determine the types of variables should be more objective and cost-effective than the current methods. In this study, we proposed a method which has not previously been available in the literature to solve this problem. We took into account of the copula function, which expresses the relation between dependent random variables in statistics, and the local dependence function, which measures the point dependency of variables. Our basic assumption is that the distribution of the variables is known in advance or can be determined.

In the simulation study, we showed how the algorithm of our method separates a bulk of synthetic variables as input and output variables. In the real data application, we selected a set of variables which was already defined as the input and output in the previous analysis of the same data in the literature. According to the new process, the suggested input/output variables were almost consistent with those of other studies [2]. Also efficiencies were the same for effective DMUs and gave less efficiency score for the ineffective DMU as seen in Table 5.

The results showed that this method has several advantages. If there is no pre-information, input/output variables can be determined objectively. There is no need for expert opinion, so it is a cost-effective method. However, this method depends on the variables which have known distributions and sometimes it becomes a disadvantage. Although there is Archimedean copula family in the copula theory, we preferred a basic copula function (FGM) in this study because Archimedean copulas use a generator function [21] and there is only one study [22] for Local Dependence Function related with Archimedean copulas and its functional structure is not suitable with

our algorithm.

As a conclusion, this is a new kind of approach which has a theoretical structure, to identify the input/output variables and it can be improved for future studies and adapted to data which have different distribution functions.

### Acknowledgments

The authors would like to thank the Editors and anonymous reviewers for their valuable comments.

### References

- [1] Charnes A., Cooper W.W. and Rhodes E., Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*, 2(6), 429-444 (1978).
- [2] Yoluk, M., Hastane Performasının Veri Zarflama Analizi (VZA) Yöntemi İle Değerlendirilmesi. Thesis (MSc). Atılım University (2010).
- [3] Morita, H. and Avkiran, N.K., Selecting Inputs and Outputs in Data Envelopment Analysis by Designing Statistical Experiments. *Journal of Operational Research Society of Japan*, 52, 163-173 (2009).
- [4] Nelsen, R.B., *An Introduction to Copulas*. Springer, New York (1999).
- [5] Bairamov, I. and Kotz, S., On Local Dependence Function for Multivariate Distributions. *New Trends in Prob. and Stat.*, 5, 27-44 (2000).
- [6] Bairamov, I., Kotz, S. and Kozubowski, T.J., A New Measure of Linear Local Dependence. *Statistics*, 37(3), 243-258 (2003).
- [7] Ramanathan, R., *An Introduction to Data Envelopment Analysis*. Sage Publications, New Delhi (2003).
- [8] Luo, Y., Bi, G. and Liang, L., Input/Output Indicator Selection for DEA Efficiency Evaluation: An Empirical Study of Chinese Commercial Banks. *Expert System with Applications*, 39, 1118-1123 (2012).
- [9] Pastor, J.T., Ruiz J.L. and Sirvent, I., A Statistical Test for Nested Radial Dea Models. *INFORMS*, 50(4), 728-735 (2002).
- [10] Jenkins, L. and Anderson, M., A multivariate statistical approach to reducing the number of variables in data envelopment analysis. *European Journal of Operational Research*, 147, 51-61 (2003).
- [11] Ruggiero, J., Impact Assessment of Input Omission on DEA. *International Journal of Information Technology and Decision Making*, 4(3), 359-368, (2005).
- [12] Edirisinghe, N.C.P. and Zhang, X., Generalized DEA Model of Fundamental Analysis and Its Application to Portfolio Optimization. *Journal of Banking and Finance*, 31, 311-335, (2007).
- [13] Madhanagopal, R. and Chandrasekaran, R., Selecting Appropriate Variables for DEA Using Genetic Algorithm (GA) Search Procedure. *International Journal of Data Envelopment Analysis and Operations Research*, 1(2), 28-33 (2014).
- [14] 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).
- [15] Fare, R. and Grosskopf, S., and Lovell, C.A.K., *Production Frontiers*. Cambridge University Press (1993).
- [16] Martić, M.M., Novaković, M.S. and Baggia, A., Data Envelopment Analysis - Basic Models and their Utilization, Organizacija. *Journal of Management, Informatics and Human Resources*, 42(2), 37-43 (2009).
- [17] Cooper, W.W., Seiford, L.M. and Zhu, J., *Handbook on Data Envelopment Analysis*. Springer New York (2011).
- [18] Peña, V.H. and Ibragimov, R., Sharakhmetov, S., Characterizations of Joint Distributions, Copula Information, Dependence and Decoupling, with Applications to Time Series. *IMS Lecture Notes-Monograph Series*, 2nd Lehmann Symposium-Optimality, 49, 183-209 (2006).
- [19] Bairamov, I. and Kotz, S., Dependence Structure and Symmetry of Huang-Kotz FGM Distributions and Their Extensions. *Metrika*, 56, 55-72 (2002).
- [20] Bairamov, I., Kotz, S. and Bekci, M., New Generalized Farlie-Gumbel-Morgenstern Distributions and Concomitants of Order Statistics. *Journal of Applied Statistics*, 28(5), 521-536 (2001).
- [21] Najjari, V., *Yeni Arşimedyan Kapula Aileleri ve Finans Alanında Bir Uygulama*. Thesis (PhD). Gazi University (2014).
- [22] Vandenhende, F., and Lambert, P., Local dependence estimation using semiparametric Archimedean copulas. *The Canadian Journal of Statistics*, 33(3), 377-388, (2005).

*Olçay Alpay is an Assistant Professor, lecturing at Sinop University in Turkey since 2011. She studied for Msc. and Ph.D in Statistics at Ege University. Her expertise is on Theory of Statistics. Her areas of interest include applied statistics and DEA.*

*Elvan Aktürk Hayat is an Assistant Professor in Department of Econometrics, Aydın Faculty of Economics, Adnan Menderes University, Turkey. She received her Ph.D degree from the Department of Statistics, Ege University, Turkey in 2012. Her areas of interest include applied statistics, data envelopment analysis and data mining.*



An International Journal of Optimization and Control: Theories & Applications (<http://ijocta.balikesir.edu.tr>)



This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit <http://creativecommons.org/licenses/by/4.0/>.