



Contents lists available at FOMJ

# Fuzzy Optimization and Modelling

Journal homepage: <http://fomj.qaemiau.ac.ir/>

**Paper Type: Research Paper**

## Comparison of the Dependence Structures of Stochastic Copula-DEA Model

**Sima Balak<sup>a</sup>, Mohammad Hassan Behzadi<sup>a,\*</sup>, Ali Nazari<sup>b</sup>**

<sup>a</sup> Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran

<sup>b</sup> Department of Statistics, Abhar Branch, Islamic Azad University, Abhar, Iran

### ARTICLE INFO

#### Article history:

Received 21 October 2021

Revised 8 November 2021

Accepted 9 November 2021

Available online 9 November 2021

#### Keywords:

Stochastic Data Envelopment Analysis

Copula Theory

Copula-SCCR

### ABSTRACT

The stochastic data envelopment analysis (SDEA) is based on treating input values and output values as random variables. Typically, in research related to stochastic data envelopment analysis, variables are assumed to be independent of one another. It is likely that there will be a dependency structure between variables. We investigated the dependence structure between input variables and between output variables. To estimate this dependence structure, we used the Copula approach. Therefore, we have proposed a stochastic DEA model with a dependence structure called Copula-SDEA in this study and evaluated this model using both input-oriented and output-oriented models. We evaluated the proposed models using real data from 10 car companies. The implementation of the proposed model showed that, different results can be drawn when taking into account the dependence structure between stochastic inputs and (or) outputs. Additionally, a comparison of Copula-SDEA models with the SDEA model revealed that the evaluation of the efficiency of DMUs with Copula-SDEA models differed significantly from the SDEA model. Moreover, the results indicate that in both input- and output-oriented models, considering the dependence structure between inputs is more important than considering the dependence structure between outputs.

## 1. Introduction

The Data Envelopment Analysis (DEA), introduced by Charnes et al. [3], is a linear programming method for computing the relative efficiency of a set of homogeneous units, called Decision Making Units (DMUs). DMUs are considered efficient if their efficiency score is equal to one, otherwise, they are considered inefficient. Through the use of a ratio of the weighted sum of outputs to the weighted sum of inputs, DEA generalizes the intuitive single-input single-output ratio efficiency measurement into a multiple-input multiple-output model [25]. DEA is an optimization approach that differs from Markowitz by focusing on the objective function of a multi constrained linear programming model, as opposed to the mean and variance parameters derived from Markowitz [17, 36].

\* Corresponding author

E-mail address: behzadi@srbiau.ac.ir (Mohammad Hassan Behzadi)

In classic DEA models, the inputs and outputs are assumed to be deterministic. In real-world applications of DEA models, variables are, however, often vague or random. Data uncertainty is the main challenge in the application of DEA models. Peykani et al. [28] categorize uncertain DEAs into five categories: Bootstrap DEA, Robust DEA, Imprecise DEA, Fuzzy DEA and Stochastic DEA. Fuzzy DEA and stochastic DEA have been widely used among these methods.

Tavan et al. [34] proposed the fuzzy stochastic data envelopment analysis with an application to base realignment and closure. Nasserri et al. [22] proposed the fuzzy stochastic input-oriented primal data envelopment analysis models with application to insurance industry. A fuzzy stochastic DEA model with undesirable outputs was proposed in [8, 23]. Pekani et al. [26] conducted a structured literature review of the fuzzy chance-constrained data envelopment analysis (FCCDEA) approach and proposed how to use it for stock valuation and portfolio selection when data is ambiguous in [27]. In [30], a new method is proposed for solving DEA models with intuitionistic fuzzy data. Using an alphabetical approach, the model is transformed into a linear programming problem with an intuitive fuzzy objective function. The fuzzy chance-constrained data envelopment analysis (FCCDEA) approach is presented in [29] for stock evaluation and portfolio selection under data ambiguity.

In DEA, a probability distribution can be used to model input/output uncertainty by modeling their inherent randomness. Stochastic DEA models are constructed by employing these distributions [7, 21, 37]. Banker et al. [2] integrated statistical elements into DEA and developed a non-parametric method. A stochastic DEA model using chance-constrained programming was proposed by [16, 32] to address the randomness of input and output data. The DEA models used in [10, 11, 14, 15, 18, 31] were proposed for stochastic data with normal distributions for the input and output variables. The stochastic DEA model of [39] uses a reliability constraint to maximize the lower bound of an entity's efficiency score with some predetermined probability. Kheirollahi et al. [13] developed a stochastic DEA with a chance-constrained programming approach that uses a relaxed combination of inputs to model congestion. The most common stochastic frontier model with an error structure was proposed by [12] on the basis of a convolution of the normal and half-normal distributions. In [24], Nazari and Behzadi proposed a stochastic DEA with skew-normal (SN) variables as the asymptotic distributions of a set of random variables.

Fuzzy stochastic DEA is another concept that has been proposed by many researchers [1, 8, 22, 23, 34, 35, 38]. Assumptions are made in all the SDEA models developed by the researchers that there is no dependency between input variables and between output variables. However, there may be dependencies between variables in real-world applications. A dependence structure can affect the output of an SDEA model. Our study assumes that the random variables follow a normal distribution and considers the dependence structure between them in the SDEA model. In order to calculate the dependence structure between variables, we use the Copula approach. A copula function is used to represent multivariate dependency. Copula theory is introduced following Sklar's theorem [33], which states that multivariate probability density function can be expressed as the product of its marginal density function and the copula density function, which represents the dependency structure among random variables. There are several families of Copula functions. Because the variables follow a normal distribution, we use the Gaussian copula to estimate the dependence structures between variables. In order to discover whether it is more important to consider the dependence structure among inputs or the dependence structure among outputs, we examine two input- and output-oriented SDEA models. As a result, we have presented input-oriented Copula-SDEA and output-oriented Copula-SDEA models in this research.

The rest of the paper is organized as follows. In Section 2, the preliminaries including the DEA concepts and its stochastic model, along with Copula theory and Copulas family are described. In Section 3, we formulate the Copula-SCCR models, and in Section 4, we evaluate them with an example. Finally, the results of the proposed models are discussed in Section 5.

## 2. Stochastic DEA

We assume  $x_j = (x_{1j}, \dots, x_{mj})$  and  $y_j = (y_{1j}, \dots, y_{sj})$  are random inputs and outputs that are related to DMU<sub>*j*</sub>,  $j = 1, \dots, n$ . It has been assumed that these components are normally distributed, i.e.  $X_{ij} \sim N(\mu_{ij}, \sigma^2_{ij})$  and  $Y_{rj} \sim N(\eta_{rj}, \upsilon^2_{rj})$ . We also assume that different DMUs' *i*th input and *r*th output are independent. Using stochastic data, an input-oriented CCR model (SCCR) can be constructed as follows:

$$\begin{aligned} \theta^*(\alpha) &= \min \theta \\ \text{s.t.} \\ P\left(\sum_{j=1}^n \lambda_j X_{ij} \leq \theta X_{io}\right) &\geq 1 - \alpha, \quad i = 1, 2, \dots, m, \\ P\left(\sum_{j=1}^n \lambda_j Y_{rj} \geq Y_{ro}\right) &\geq 1 - \alpha, \quad r = 1, 2, \dots, s, \\ \lambda_j &\geq 0, \quad j = 1, \dots, n. \end{aligned} \tag{1}$$

where  $\alpha$  is the level of error between 0 and 1, which has been predetermined. By using the external slack variables in the model (1), we convert the inequality constraints to equality constraints according to Khodabakhshi and Asgharian [14]. Therefore, we can express the stochastic version of the model (1) as follows:

$$\begin{aligned} \min \theta - \varepsilon \left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right) \\ \text{s.t.} \\ P\left(\sum_{j=1}^n \lambda_j X_{ij} - \theta X_{io} \leq -s_i^-\right) &= 1 - \alpha, \quad i = 1, \dots, m, \\ P\left(\sum_{j=1}^n \lambda_j Y_{rj} - Y_{ro} \geq s_r^+\right) &= 1 - \alpha, \quad r = 1, \dots, s, \\ \lambda_j, s_i^-, s_r^+ &\geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s. \end{aligned} \tag{2}$$

**Definition 1.** (*Efficiency*). The DMU<sub>o</sub> is Pareto efficient if and only if in optimal solution [3]:

1.  $\theta^* = 1$ .
2.  $s_i^- = s_r^+ = 0, \quad \forall i, r$ .

In order to convert model (2) to a deterministic form, given that  $X_{ij}$  and  $Y_{rj}$  have a normal distribution, we use the expected value and input and output variances as described by Dibachi et al. [6]:

$$\begin{aligned} E\left(\sum_{j=1}^n \lambda_j X_{ij} - \theta X_{io}\right) &= \sum_{j=1}^n \lambda_j \mu_{ij} - \theta \mu_{io}, \\ E\left(\sum_{j=1}^n \lambda_j Y_{rj} - Y_{ro}\right) &= \sum_{j=1}^n \lambda_j \eta_{rj} - \eta_{ro}. \end{aligned} \tag{3}$$

$$\begin{aligned}\sigma_i^2(\lambda, \theta) &= \text{var} \left( \sum_{j=1}^n \lambda_j X_{ij} - \theta X_{io} \right), \\ &= \sum_{k=1}^n \sum_{j=1}^n \lambda_j \lambda_k \text{cov}(X_{ik}, X_{ij}) - 2\theta \sum_{j=1}^n \lambda_j \text{cov}(X_{ij}, X_{io}) + \theta^2 \sigma_{io}^2.\end{aligned}\quad (4)$$

$$\begin{aligned}\nu_r^2(\lambda) &= \text{var} \left( \sum_{j=1}^n \lambda_j Y_{rj} - Y_{ro} \right) \\ &= \sum_{k=1}^n \sum_{j=1}^n \lambda_j \lambda_k \text{cov}(Y_{rk}, Y_{rj}) - 2 \sum_{j=1}^n \lambda_j \text{cov}(Y_{rj}, Y_{ro}) + \nu_{ro}^2.\end{aligned}\quad (5)$$

Accordingly, the deterministic form of model (2) can be expressed as follows:

$$\begin{aligned}\min \theta - \varepsilon &\left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right) \\ \text{s.t.} \\ \sum_{j=1}^n \lambda_j \mu_{ij} + s_i^- - \Phi^{-1}(\alpha) \sigma_i(\lambda, \theta) &= \theta \mu_{io}, \quad i = 1, 2, \dots, m, \\ \sum_{j=1}^n \lambda_j \eta_{rj} - s_r^+ + \Phi^{-1}(\alpha) \nu_r(\lambda) &= \eta_{ro}, \quad r = 1, 2, \dots, s, \\ \lambda_j, s_i^-, s_r^+ &\geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s.\end{aligned}\quad (6)$$

where,  $\Phi$  is the cumulative distribution function (CDF) of the normal distribution.

According to the same procedure, output-oriented SCCR model is derived as follows:

$$\begin{aligned}\max \phi + \varepsilon &\left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right) \\ \text{s.t.} \\ \sum_{j=1}^n \lambda_j \mu_{ij} + s_i^- - \Phi^{-1}(\alpha) \sigma_i(\lambda, \theta) &= \mu_{io}, \quad i = 1, 2, \dots, m, \\ \sum_{j=1}^n \lambda_j \eta_{rj} - s_r^+ + \Phi^{-1}(\alpha) \nu_r(\lambda) &= \phi \eta_{ro}, \quad r = 1, 2, \dots, s, \\ \lambda_j, s_i^-, s_r^+ &\geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s.\end{aligned}\quad (7)$$

### 3. Dependence structure and copula

The dependence between random variables is of great significance since it may indicate important statistical or causal relationships in real-world systems. Therefore, it is crucial to estimate such dependencies between variables. The Copula theory is one of the best approaches to estimating the dependence structure. A Copula ( $C$ ) is multivariate cumulative distribution function with uniform univariate margins on  $[0, 1]$ , and is defined as follows [20]:

**Definition 2.** (Copula). Given  $d \geq 2$  random variables  $\mathbf{X} = \{X_1, \dots, X_d\} \in \mathfrak{R}^d$ . Consider  $\{u_i = F_i(x_i), i = 1, \dots, d\}$  to be

a marginal distribution. The  $d$ -dimensional copula  $C : I^d \rightarrow I$  ( $I = [0,1]$ ) of  $\mathbf{X}$  has the following properties:

- If any of the arguments is zero, the copula is zero;  $C(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0$ .
- If one argument is  $u$  and all others are 1, then the copula equals  $u$ ;  $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ .

A copula’s application is theoretically based on Sklar’s theorem.

**Theorem 1.** (Sklar’s Theorem). Given  $d \geq 2$  random variables  $\mathbf{X} = \{X_1, \dots, X_d\} \in \mathfrak{R}^d$ , its cumulative distribution function (CDF)  $\mathbf{F}(\mathbf{x})$  can be written as follow:

$$\mathbf{F}(\mathbf{x}) = C\{u_1, \dots, u_d\}. \tag{8}$$

where  $\{u_i\} \in (0,1)$  are marginal distribution functions of  $\mathbf{X}$ . if  $\{F_i\}$  are continuous, then  $C$  is unique.

Researchers have developed several families of copulas. The Gaussian and the Archimedean copulas are the most widely used Copula functions. Archimedean copulas are a class of associative copulas, which include several families of Copula functions, such as the Frank, Joe, Gumbel, and Clayton functions. Due to the normal distribution of the stochastic variables, the Gaussian copula was used in this study. Following is a brief explanation of this copula.

### 3.1. Gaussian Copula

A Gaussian copula is a distribution over a unit hyper cube  $[0,1]^d$ . Bivariate Gaussian copulas can be constructed using a bivariate normal distribution with unit variances, zero means, and correlation  $\rho$ . The Gaussian copula can be expressed as follows for a given correlation parameter  $\rho$ :

$$C(u_1, u_2; \rho) = \Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho), \quad -1 \leq \rho \leq 1. \tag{9}$$

where  $\Phi^{-1}$  is the inverse cumulative distribution function of a standard normal and  $\Phi_\rho$  denote the joint cumulative distribution function of a bivariate normal distribution with the mean vector zero and covariance matrix equal to the correlation  $\rho$ . The corresponding copula density can be expressed as [20],

$$C(u_1, u_2; \rho) = \frac{1}{\phi(x_1)\phi(x_2)} \frac{1}{\sqrt{1-\rho^2}} \times \exp\left\{\frac{\rho^2(x_1^2 + x_2^2) - 2\rho x_1 x_2}{2(1-\rho^2)}\right\}. \tag{10}$$

where  $x_1 = \Phi^{-1}(u_1)$  and  $x_2 = \Phi^{-1}(u_2)$ . Kendall’s  $\tau$  of Gaussian copula is as follow,

$$\tau = \frac{2}{\pi} \arcsin(\rho). \tag{11}$$

### 3.2. Estimating Copulas

Several methods can be used to estimate copulas, including maximum likelihood estimation (MLE), Pseudo-MLE, and Moment-matching. We estimate Copulas using maximum likelihood estimation (MLE). Based on Hofert et al. [9], given realizations  $u_i, i \in \{1, \dots, d\}$  of a random sample  $U_i, i \in \{1, \dots, d\}$ , and the copula  $C$ , the log-likelihood is defined by,

$$\ln(u_1, \dots, u_d; \kappa) = \sum_{i=1}^d \ln(u_i; \kappa). \tag{12}$$

The ML estimator  $\hat{\kappa}_n$  is obtained by numerically maximizing (12) with respect to  $\kappa$ :

$$\hat{\kappa}_n = \arg \sup \ln(u_1, \dots, u_d; \kappa). \tag{13}$$

#### 4. Formulation of the SCCR models with dependence structure

There may be a dependence structure between random variables, as was mentioned earlier. Analysis of SDEA models without taking this dependence structure into account can give incorrect results and result in incorrect decisions. We consider separately the dependence structure between input variables and the dependence structure between output variables. To evaluate the impact of these dependence structures, we presented two models: 1- input-oriented SCCR model with dependence structure, and 2- output-oriented SCCR model with dependence structure. Since we estimate the dependence structure using copula, we call these models input-oriented copula-SCCR (IO-Copula-SCCR) and output-oriented copula-SCCR (OO-Copula-SCCR).

Let us consider  $m$  random variables with normal distributions  $x_1, \dots, x_m$ . The pairwise dependence values of these variables are represented by  $\tau_{x_i x_j}, i = 1, \dots, m; j = 1, \dots, m; i \neq j$  ( $\tau_{x_i x_j} = \tau_{x_j x_i}$ ). We define the new variables as follows:

$$x_{\text{new}i} = x_i + \sum_{i \neq j}^m \tau_{x_i x_j} x_j, \quad \text{for } i = 1, 2, \dots, m. \quad (14)$$

In fact, the values of the variables are added according to the degree of dependence of its distribution on other variables. In general, it can be said that the new variable is equal to its sum with the weighted sum of the other variables, where the weights are equal to the values of the dependence of the distribution of the variables.

**Theorem 2.** (*Dependence structure*). The dependence structure causes changes in the mean ( $\mu$ ) and variance ( $\sigma^2$ ) of any random variable with a normal distribution.

**Proof.** Consider a random variable  $X_i$  with a normal distribution  $X_i \sim N(\mu_i, \sigma_i^2)$ . Linear combinations obtained from  $X_i$  have a normal distribution as well. Therefore, taking Eq. (14) into account, we have:

$$\mu = E(x_{\text{new}i}) = E\left(x_i + \sum_{i \neq j}^m \tau_{x_i x_j} x_j\right) \Rightarrow \mu = \mu_i + \sum_{i \neq j}^m \tau_{x_i x_j} \mu_j \quad (15)$$

$$\begin{aligned} \sigma(\tau)^2 &= \text{var}(x_{\text{new}i}) = \text{var}\left(x_i + \sum_{i \neq j}^m \tau_{x_i x_j} x_j\right), \\ \Rightarrow \sigma(\tau)^2 &= \text{var}(x_i) + \text{var}\left(\sum_{i \neq j}^m \tau_{x_i x_j} x_j\right) + 2 \text{cov}\left(x_i, \sum_{i \neq j}^m \tau_{x_i x_j} x_j\right), \\ \Rightarrow \sigma(\tau)^2 &= \sigma_i + \sum_{i \neq j}^m \sum_{k=1}^m \tau_{x_i x_j} \tau_{x_i x_k} \text{cov}(x_j, x_k) + 2 \text{cov}\left(x_i, \sum_{i \neq j}^m \tau_{x_i x_j} x_j\right). \end{aligned} \quad (16)$$

This completes the proof.  $\square$

##### 4.1. IO-Copula-SCCR model with input dependence structure

Let the distributions of input variables be dependent, but let the output variables be independent. The IO-Copula-SCCR model is expressed as follows, taking into account the model (2) and the dependence between the distributions of input variables:

$$\theta^*(\alpha) = \min \theta$$

s.t.

$$P\left(\sum_{j=1}^n \lambda_j \left(X_{ij} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_j} X_{(k)_j}\right) \leq \theta \left(X_{io} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_o} X_{(k)_o}\right)\right) \geq 1 - \alpha, \quad i = 1, \dots, m, \tag{17}$$

$$P\left(\sum_{j=1}^n \lambda_j Y_{rj} \geq Y_{ro}\right) \geq 1 - \alpha, \quad r = 1, 2, \dots, s,$$

$$\lambda_j \geq 0, \quad j = 1, \dots, n.$$

In order to convert inequality constraints into equality constraints in model (17), there exists  $s_i^- > 0$ , and  $s_r^+ > 0$  such that:

$$P\left(\sum_{j=1}^n \lambda_j \left(X_{ij} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_j} X_{(k)_j}\right) - \theta \left(X_{io} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_o} X_{(k)_o}\right) \leq -s_i^-\right) = 1 - \alpha. \tag{18}$$

$$P\left(\sum_{j=1}^n \lambda_j Y_{rj} - Y_{ro} \geq s_r^+\right) = 1 - \alpha. \tag{19}$$

Therefore, model (17) can be rewritten as follows:

$$\theta^*(\alpha) = \min \theta - \varepsilon \left(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+\right),$$

s.t.

$$P\left(\sum_{j=1}^n \lambda_j \left(X_{ij} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_j} X_{(k)_j}\right) - \theta \left(X_{io} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_o} X_{(k)_o}\right) \leq -s_i^-\right) = 1 - \alpha, \quad i = 1, \dots, m, \tag{20}$$

$$P\left(\sum_{j=1}^n \lambda_j Y_{rj} - Y_{ro} \geq s_r^+\right) = 1 - \alpha, \quad r = 1, \dots, s,$$

$$\lambda_j, s_i^-, s_r^+ \geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s.$$

**Definition 3.** (Stochastic Efficiency). DMU<sub>o</sub> is Pareto efficient for every significance level  $\alpha$ , if and only if in the optimal solution,

1.  $\theta^* = 1$ .
2.  $s_i^- = s_r^+ = 0, \forall i, r$ .

In order to obtain the deterministic form of the model (20), we consider  $h_i$  as follows:

$$h_i = \sum_{j=1}^n \lambda_j \left(X_{ij} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_j} X_{(k)_j}\right) - \theta \left(X_{io} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_o} X_{(k)_o}\right). \tag{21}$$

Since  $X_{ij}$  and  $Y_{rj}$  are assumed to have normal distributions,  $h_i$  will also have a normal distribution  $h_i \sim N(\bar{\mu}_i, \bar{\sigma}_i(\lambda, \tau, \theta)^2)$ . Where,

$$\bar{\mu}_i = E(h_i) = \sum_{j=1}^n \lambda_j \left( \mu_{ij} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_j} \mu_{(k)_j} \right) - \theta \left( \mu_{io} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_o} \mu_{(k)_o} \right). \tag{22}$$

$$\begin{aligned} \bar{\sigma}_i(\lambda, \tau, \theta)^2 &= \text{var} \left( \sum_{j=1}^n \lambda_j A_j - \theta B_o \right), \\ &= \text{var} \left( \sum_{j=1}^n \lambda_j A_j \right) + \text{var}(\theta B_o) - 2 \text{cov} \left( \sum_{j=1}^n \lambda_j A_j, \theta B_o \right), \\ &= \sum_{j=1}^n \sum_{k=1}^n \lambda_j \lambda_k \text{cov}(A_j, A_k) + \theta^2 \text{var}(B_o) - 2\theta \sum_{j=1}^n \lambda_j \text{cov}(A_j, B_o). \end{aligned} \tag{23}$$

where,  $A_j = \left( X_i + \sum_{i \neq k}^m \tau_{x_i, x_k} X_k \right)_j$ , and  $B_o = \left( X_i + \sum_{i \neq k}^m \tau_{x_i, x_k} X_k \right)_o$ .

Given the stochastic variable  $h_i$ , the first inequality of model (20) can be rewritten in the following form:

$$P(h_i \leq -s_i^-) = 1 - \alpha \Rightarrow P\left( \frac{h_i - \bar{\mu}_i}{\bar{\sigma}_i(\lambda, \tau, \theta)} \leq \frac{-s_i^- - \bar{\mu}_i}{\bar{\sigma}_i(\lambda, \tau, \theta)} \right) = 1 - \alpha. \tag{24}$$

Taking  $Z_i = \frac{h_i - \bar{\mu}_i}{\bar{\sigma}_i(\lambda, \tau, \theta)}$  and  $Z_i$  to be normally distributed, we get:

$$P\left( Z_i \leq \frac{-s_i^- - \bar{\mu}_i}{\bar{\sigma}_i(\lambda, \tau, \theta)} \right) = 1 - \alpha \Rightarrow P\left( Z_i \leq \frac{s_i^- + \bar{\mu}_i}{\bar{\sigma}_i(\lambda, \tau, \theta)} \right) = \alpha \Rightarrow \Phi\left( \frac{s_i^- + \bar{\mu}_i}{\bar{\sigma}_i(\lambda, \tau, \theta)} \right) = \alpha. \tag{25}$$

where  $\Phi$  is the standard normal cumulative distribution function. Thus:

$$\frac{s_i^- + \bar{\mu}_i}{\bar{\sigma}_i(\lambda, \tau, \theta)} = \Phi^{-1}(\alpha) \Rightarrow \bar{\mu}_i + s_i^- - \Phi^{-1}(\alpha) \bar{\sigma}_i(\lambda, \tau, \theta) = 0. \tag{26}$$

As a result, the probability constraint for the  $i$ -th input in model (20) can be expressed in the following deterministic form:

$$\sum_{j=1}^n \lambda_j \bar{\mu}_{ij} + s_i^- - \Phi^{-1}(\alpha) \bar{\sigma}_i(\lambda, \tau, \theta) = \theta \bar{\mu}_{io} \tag{27}$$

where  $\bar{\mu}_{ij} = \mu_{ij} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_j} \mu_{(k)_j}$ , and  $\bar{\mu}_{io} = \mu_{io} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_o} \mu_{(k)_o}$ .

Similarly, the deterministic form of the probability constraint for the  $r$ -th output in model (20) is given by:

$$\sum_{j=1}^n \lambda_j \eta_{rj} - s_r^+ + \Phi^{-1}(\alpha) \nu_r(\lambda) = \eta_{ro}. \tag{28}$$

where

$$\eta_{rj} = E(y_{rj}). \tag{29}$$

$$\nu_r(\lambda)^2 = \sum_{j=1}^n \sum_{k=1}^n \lambda_j \lambda_k \text{cov}(y_{rj}, y_{rk}) + \text{var}(y_{ro}) - 2 \sum_{j=1}^n \lambda_j \text{cov}(y_{rj}, y_{ro}). \tag{30}$$



Accordingly, the deterministic form of the IO-Copula-SCCR model (20) with input dependence structure is as follows:

$$\begin{aligned} \min \theta - \varepsilon & \left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right), \\ \text{s.t.} & \\ & \sum_{j=1}^n \lambda_j \bar{\mu}_{ij} + s_i^- - \Phi^{-1}(\alpha) \bar{\sigma}_i(\lambda, \tau, \theta) = \theta \bar{\mu}_{io}, \quad i = 1, 2, \dots, m, \\ & \sum_{j=1}^n \lambda_j \eta_{rj} - s_r^+ + \Phi^{-1}(\alpha) \nu_r(\lambda) = \eta_{ro}, \quad r = 1, 2, \dots, s, \\ & \lambda_j, s_i^-, s_r^+ \geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s. \end{aligned} \tag{31}$$

**Theorem 3.** For every level  $\alpha$ , model (31) is feasible.

**Proof.** Let  $\lambda_o = 1, \theta = 1, \lambda_j = 0, j \neq o$  and also  $s_i^- = s_r^+ = 0, i = 1, 2, \dots, m, r = 1, 2, \dots, s$ . □

Clearly, if we do not take the dependence between the input distributions into account ( $\tau = 0$ ) in the model (31), then it results in a SCCR model.

#### 4.2. IO-Copula-SCCR model with output dependence structure

Let the distributions of output variables be dependent, but let the input variables be independent. The IO-Copula-SCCR model is expressed as follows, taking into account the model (2) and the dependence between the distributions of output variables:

$$\begin{aligned} \theta^*(\alpha) &= \min \theta \\ \text{s.t.} & \\ & P \left( \sum_{j=1}^n \lambda_j X_{ij} \leq \theta X_{io} \right) \geq 1 - \alpha, \quad i = 1, 2, \dots, m, \\ & P \left( \sum_{j=1}^n \lambda_j \left( Y_{rj} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_j} Y_{(g)_j} \right) \geq \left( Y_{ro} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_o} Y_{(g)_o} \right) \right) \geq 1 - \alpha, \quad r = 1, 2, \dots, s, \\ & \lambda_j \geq 0, \quad j = 1, 2, \dots, n. \end{aligned} \tag{32}$$

In order to convert inequality constraints into equality constraints in model (32), there exists  $s_i^- > 0$ , and  $s_r^+ > 0$  such that:

$$P \left( \sum_{j=1}^n \lambda_j \left( Y_{rj} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_j} Y_{(g)_j} \right) - \left( Y_{ro} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_o} Y_{(g)_o} \right) \geq s_r^+ \right) = 1 - \alpha. \tag{33}$$

$$P \left( \sum_{j=1}^n \lambda_j X_{ij} - \theta X_{io} \leq -s_i^- \right) = 1 - \alpha. \tag{34}$$

Therefore, model (32) can be rewritten as follows:

$$\min \theta - \varepsilon \left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right),$$

s.t.

$$P \left( \sum_{j=1}^n \lambda_j X_{ij} - \theta X_{io} \leq -s_i^- \right) = 1 - \alpha, \quad i = 1, 2, \dots, m, \quad (35)$$

$$P \left( \sum_{j=1}^n \lambda_j \left( Y_{rj} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_j} Y_{(g)_j} \right) - \left( Y_{ro} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_o} Y_{(g)_o} \right) \geq s_r^+ \right) = 1 - \alpha, \quad r = 1, 2, \dots, s,$$

$$\lambda_j, s_i^-, s_r^+ \geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s.$$

According to definition 3, DMUo is Pareto efficient at any significance level  $\alpha$  if and only if:

1.  $\theta^* = 1$ .
2.  $s_i^- = s_r^+ = 0, \quad \forall i, r$ .

Using the procedure outlined in the previous section, the first and second constraints on model (36) take the following forms, respectively:

$$\sum_{j=1}^n \lambda_j \mu_{ij} + s_i^- - \Phi^{-1}(\alpha) \sigma_i(\lambda, \theta) = \theta \mu_{io}. \quad (36)$$

$$\sum_{j=1}^n \lambda_j \bar{\eta}_{rj} - s_r^+ + \Phi^{-1}(\alpha) \bar{\nu}_r(\lambda, \tau) = \bar{\eta}_{ro}. \quad (37)$$

where,

$$\mu_{ij} = E(X_{ij}). \quad (38)$$

$$\sigma_i(\lambda, \theta)^2 = \sum_{j=1}^n \sum_{k=1}^n \lambda_j \lambda_k \text{cov}(X_{ij}, X_{ik}) + \theta^2 \text{var}(X_{io}) - 2 \sum_{j=1}^n \lambda_j \text{cov}(X_{ij}, X_{io}). \quad (39)$$

$$\bar{\eta}_{rj} = \eta_{rj} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_j} \eta_{(g)_j}, \quad (40)$$

$$\bar{\eta}_{ro} = \eta_{ro} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_o} \eta_{(g)_o}.$$

$$\begin{aligned} \bar{\nu}_r(\lambda, \tau)^2 &= \text{var} \left( \sum_{j=1}^n \lambda_j A'_j - B'_o \right), \\ &= \text{var} \left( \sum_{j=1}^n \lambda_j A'_j \right) + \text{var}(B'_o) - 2 \text{cov} \left( \sum_{j=1}^n \lambda_j A'_j, B'_o \right), \\ &= \sum_{j=1}^n \sum_{k=1}^n \lambda_j \lambda_k \text{cov}(A'_j, A'_k) + \text{var}(B'_o) - 2 \sum_{j=1}^n \lambda_j \text{cov}(A'_j, B'_o). \end{aligned} \quad (41)$$

$$\text{where, } A'_j = \left( Y_r + \sum_{r \neq g}^s \tau_{Y_r Y_g} Y_g \right)_j \text{ and } B'_o = \left( Y_r + \sum_{r \neq g}^s \tau_{Y_r Y_g} Y_g \right)_o.$$

Accordingly, the deterministic IO-Copula-SCCR model with output dependence is as follows

$$\begin{aligned}
 & \min \theta - \varepsilon \left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right), \\
 & \text{s.t.} \\
 & \sum_{j=1}^n \lambda_j \mu_{ij} + s_i^- - \Phi^{-1}(\alpha) \sigma_i(\lambda, \theta) = \theta \mu_{io}, \quad i = 1, 2, \dots, m, \\
 & \sum_{j=1}^n \lambda_j \bar{\eta}_{rj} - s_r^+ + \Phi^{-1}(\alpha) \bar{\nu}_r(\lambda, \tau) = \bar{\eta}_{ro}, \quad r = 1, 2, \dots, s, \\
 & \lambda_j, s_i^-, s_r^+ \geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s.
 \end{aligned} \tag{42}$$

**Theorem 4.** Model (42) is feasible for every significant level  $\alpha$ .

**Proof.** Same as Theorem 2.  $\square$

**4.3. OO-Copula-SCCR model with input dependence structure**

Considering the dependence structure between input variables, we can write the output-oriented SCCR model (7) as follows:

$$\begin{aligned}
 & \phi^*(\alpha) = \max \phi \\
 & \text{s.t.} \\
 & P \left( \sum_{j=1}^n \lambda_j \left( X_{ij} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_j} X_{(k)_j} \right) \leq \left( X_{io} + \sum_{i \neq k}^m \tau_{(x_i, x_k)_o} X_{(k)_o} \right) \right) \geq 1 - \alpha, \quad i = 1, \dots, m, \\
 & P \left( \sum_{j=1}^n \lambda_j Y_{rj} \geq \phi Y_{ro} \right) \geq 1 - \alpha, \quad r = 1, 2, \dots, s, \\
 & \lambda_j \geq 0, \quad j = 1, \dots, n.
 \end{aligned} \tag{43}$$

By following the same procedure described in Section 4.1, we are able to write the deterministic form of the OO-Copula-SCCR model (43) with input dependence structure is as follows:

$$\begin{aligned}
 & \max \phi + \varepsilon \left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right) \\
 & \text{s.t.} \\
 & \sum_{j=1}^n \lambda_j \bar{\mu}_{ij} + s_i^- - \Phi^{-1}(\alpha) \bar{\sigma}_i(\lambda, \tau) = \bar{\mu}_{io}, \quad i = 1, 2, \dots, m, \\
 & \sum_{j=1}^n \lambda_j \eta_{rj} - s_r^+ + \Phi^{-1}(\alpha) \nu_r(\lambda, \phi) = \phi \eta_{ro}, \quad r = 1, 2, \dots, s, \\
 & \lambda_j, s_i^-, s_r^+ \geq 0, \quad j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s.
 \end{aligned} \tag{44}$$

**Theorem 5.** For every level  $\alpha$ , model (44) is feasible.

**Proof.** The proof is same as Theorem 2.  $\square$

#### 4.4. OO-Copula-SCCR model with output dependence structure

Considering the dependence structure between output variables, we can write the output-oriented SCCR model (7) as follows:

$$\phi^*(\alpha) = \max \phi$$

s.t.

$$\begin{aligned} P\left(\sum_{j=1}^n \lambda_j X_{ij} \leq X_{io}\right) &\geq 1 - \alpha, & i = 1, 2, \dots, m, \\ P\left(\sum_{j=1}^n \lambda_j \left(Y_{rj} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_j} Y_{(g)_j}\right)\right) &\geq \phi \left(Y_{ro} + \sum_{r \neq g}^s \tau_{(Y_r Y_g)_o} Y_{(g)_o}\right) \geq 1 - \alpha, & r = 1, 2, \dots, s, \\ \lambda_j &\geq 0, & j = 1, 2, \dots, n. \end{aligned} \quad (45)$$

By following the same procedure described in Section 4.2, we are able to write the deterministic form of the OO-Copula-SCCR model (45) with output dependence structure is as follows:

$$\max \phi + \varepsilon \left( \sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right)$$

s.t.

$$\begin{aligned} \sum_{j=1}^n \lambda_j \mu_{ij} + s_i^- - \Phi^{-1}(\alpha) \sigma_i(\lambda) &= \mu_{io}, & i = 1, 2, \dots, m, \\ \sum_{j=1}^n \lambda_j \bar{\eta}_{rj} - s_r^+ + \Phi^{-1}(\alpha) \bar{\nu}_r(\lambda, \tau, \phi) &= \phi \bar{\eta}_{ro}, & r = 1, 2, \dots, s, \\ \lambda_j, s_i^-, s_r^+ &\geq 0, & j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m, \quad r = 1, 2, \dots, s. \end{aligned} \quad (46)$$

**Theorem 6.** For every level  $\alpha$ , model (46) is feasible.

**Proof.** The proof is same as Theorem 2.  $\square$

## 5. Numerical example

In this section, we demonstrate the performance of our proposed models using real-world data. Consider 10 car companies with two stochastic inputs and two stochastic outputs. In this model, “construction costs ( $x_1$ )”, and “after-sales service costs ( $x_2$ )” are inputs and “0 to 100 km/h acceleration ( $y_1$ )”, and “maximum required power (H.P) at 6000 rpm ( $y_2$ )” are outputs. According to the goodness of fit test, the input and output data have normal distributions. The estimates are shown in Table 1.

The proposed models were solved numerically using MATLAB software. Tables 2-7 provide the relative efficiency of 10 car companies and the numerical values for shortage and surplus variables at the significant level of 0.05 for these models. There are significant differences between the results of the SCCR model and the Copula-SCCR model. Therefore, some DMUs that have been evaluated by the SCCR model as efficient, taking into consideration the structure of dependency between variables, are stochastically inefficient. Also, there are DMUs that are stochastically inefficient when analysed by the SCCR model, but are stochastically efficient when analysed by the Copula-SCCR model. It highlights the importance of considering the dependency structure and the strength of Copula's approach in estimating these dependencies.

**Table 1.** Estimated input and output parameters

DMUs	$x_1 \sim N(\mu, \sigma^2)$	$x_2 \sim N(\mu, \sigma^2)$	$y_1 \sim N(\mu, \sigma^2)$	$y_2 \sim N(\mu, \sigma^2)$
DMU1	N(1001.04, 588.71)	N(96.8643, 0.23)	N(11.42045, 0.06)	N(9.0008, 96.03)
DMU2	N(9975.08, 1263.88)	N(103.7031, 0.25)	N(11.17942, 0.30)	N(9.40659, 105.62)
DMU3	N(10017.94, 273.32)	N(98.6785, 0.02)	N(9.90928, 0.20)	N(10.14612, 17.41)
DMU4	N(10029.88, 382.65)	N(98.1702, 0.07)	N(9.55627, 0.25)	N(10.54101, 8.67)
DMU5	N(9981.57, 881.34)	N(105.022, 0.54)	N(11.93626, 0.15)	N(8.507416, 35.24)
DMU6	N(10017.25, 2697.64)	N(105.6178, 0.57)	N(11.42497, 0.21)	N(9.0617, 18.31)
DMU7	N(10025.56, 268.25)	N(90.4922, 0.02)	N(9.96474, 0.04)	N(10.39814, 730.76)
DMU8	N(9131.6, 7082547.56)	N(100.4313, 0.04)	N(9.77114, 0.18)	N(10.3222, 5.99)
DMU9	N(9986.84, 920.7)	N(99.8109, 0.74)	N(11.28765, 1.21)	N(9.25204, 46.76)
DMU10	N(19014.39, 728917699.53)	N(98.1506, 0.01)	N(10.04604, 0.05)	N(10.70154, 14.27)

**Table 2.** The efficiency and slacks of DMUs for Input-Oriented SCCR model

	DMUs	$\theta$	$s_1^-$	$s_2^-$	$s_1^+$	$s_2^+$
*	DMU1	1.000	0.000	0.000	0.000	0.000
	DMU2	0.730	0.000	0.000	0.000	0.018
	DMU3	0.998	0.000	0.000	0.000	0.000
	DMU4	1.000	0.000	0.000	0.000	0.002
*	DMU5	1.000	0.000	0.000	0.000	0.000
*	DMU6	1.000	0.000	0.000	0.000	0.000
	DMU7	0.821	0.000	0.000	0.001	0.037
*	DMU8	1.000	0.000	0.000	0.000	0.000
	DMU9	0.786	0.000	0.000	0.001	0.000
*	DMU10	1.000	0.000	0.000	0.000	0.000

**Table 3.** The efficiency and slacks of DMUs for Output-Oriented SCCR model

	DMUs	$\theta$	$s_1^-$	$s_2^-$	$s_1^+$	$s_2^+$
	DMU1	0.896	0.001	0.472	0.530	0.061
	DMU2	0.838	0.012	0.007	0.487	0.109
*	DMU3	1.000	0.000	0.000	0.000	0.000
*	DMU4	1.000	0.000	0.000	0.000	0.000
	DMU5	0.901	0.000	0.016	0.479	0.059
	DMU6	0.928	0.000	0.005	0.506	0.086
	DMU7	0.895	0.008	0.135	0.385	0.065
	DMU8	0.909	0.152	0.013	0.563	0.141
	DMU9	0.898	0.008	0.002	0.494	0.045
*	DMU10	1.000	0.000	0.000	0.000	0.000

**Table 4.** The efficiency and slacks of DMUs for IO-Copula-SCCR model with input dependence structure

	DMUs	$\theta$	$s_1^-$	$s_2^-$	$s_1^+$	$s_2^+$
*	DMU1	1.000	0.000	0.000	0.000	0.000
	DMU2	0.895	0.000	0.006	0.000	0.000
	DMU3	0.794	0.000	0.006	0.000	0.000
	DMU4	0.740	0.000	0.006	0.000	0.000
*	DMU5	1.000	0.000	0.000	0.000	0.000
*	DMU6	1.000	0.000	0.000	0.000	0.000
	DMU7	0.748	0.000	0.006	0.000	0.131
	DMU8	0.650	0.531	0.000	0.000	0.000
	DMU9	0.831	0.000	0.007	0.001	0.000
	DMU10	0.132	0.507	0.000	0.000	0.000

**Table 5.** The efficiency and slacks of DMUs for OO-Copula-SCCR model with input dependence structure

	DMUs	$\theta$	$s_1^-$	$s_2^-$	$s_1^+$	$s_2^+$
	DMU1	0.823	0.007	0.000	0.062	0.010
	DMU2	0.891	0.002	0.000	0.012	0.009
	DMU3	0.873	0.001	0.000	0.004	0.006
*	DMU4	1.000	0.000	0.000	0.000	0.000
	DMU5	0.032	0.000	0.005	0.500	0.055
	DMU6	0.022	0.000	0.008	0.500	0.055
	DMU7	0.105	0.000	0.005	0.473	0.047
	DMU8	0.177	0.008	0.000	0.357	0.035
*	DMU9	1.000	0.000	0.000	0.000	0.000
*	DMU10	1.000	0.000	0.000	0.000	0.000

**Table 6.** The efficiency and slacks of DMUs for IO-Copula-SCCR model with output dependence structure

	DMUs	$\theta$	$s_1^-$	$s_2^-$	$s_1^+$	$s_2^+$
*	DMU1	1.000	0.000	0.000	0.000	0.000
	DMU2	0.710	0.000	0.000	0.003	0.000
	DMU3	0.803	0.000	0.005	0.003	0.000
	DMU4	0.701	0.000	0.000	0.008	0.000
	DMU5	0.879	0.000	0.000	0.000	0.018
*	DMU6	1.000	0.000	0.000	0.000	0.000
	DMU7	1.000	0.000	0.000	0.000	0.207
	DMU8	0.687	0.000	0.000	0.010	0.000
	DMU9	0.784	0.000	0.001	0.000	0.004
*	DMU10	1.000	0.000	0.000	0.000	0.000

**Table 7.** The efficiency and slacks of DMUs for OO-Copula-SCCR model with output dependence structure

	DMUs	$\theta$	$s_1^-$	$s_2^-$	$s_1^+$	$s_2^+$
	DMU1	0.375	0.004	0.010	0.476	0.057
*	DMU2	1.000	0.000	0.000	0.000	0.000
	DMU3	0.132	0.011	0.004	0.470	0.059
	DMU4	0.199	0.013	0.005	0.470	0.060
*	DMU5	1.000	0.000	0.000	0.000	0.000
*	DMU6	1.000	0.000	0.000	0.000	0.000
*	DMU7	1.000	0.000	0.000	0.000	0.000
	DMU8	0.865	0.007	0.000	0.001	0.003
	DMU9	0.294	0.001	0.005	0.471	0.052
*	DMU10	1.000	0.000	0.000	0.000	0.000

The Figures 1 and 2 compare the assessed efficiencies for input- and output-oriented models, respectively. As shown in the figures, input- and output-oriented Copula-SCCR models have different estimated efficacies compared to the SCCR model. To better compare Copula-SCCR and SCCR models, Figures 3 and 4 show the differences between the estimated efficiency of input- and output-oriented Copula-SCCR models and SCCR models. As shown in Figure 3, the difference between input-oriented Copula-SCCR and SCCR models is impressive, with an average of 20.5% for input dependence case, and 11.8% for output dependence case. The Figure 4 also shows an average difference of 40% between the output-oriented Copula-SCCR and SCCR models for input dependence, and 34.7% for output dependence (Table 8). Consequently, output-oriented models have greater effects from dependencies than input-oriented models. In addition, the mean differences between the models in both input- and output-oriented cases indicate that in both input- and output-oriented cases, the Copula-SCCR model with input dependence structure differs the most from the SCCR model. A very important finding reveals that considering the dependence structure between inputs is more important than considering the dependence structure between outputs in both input- and output-oriented models.

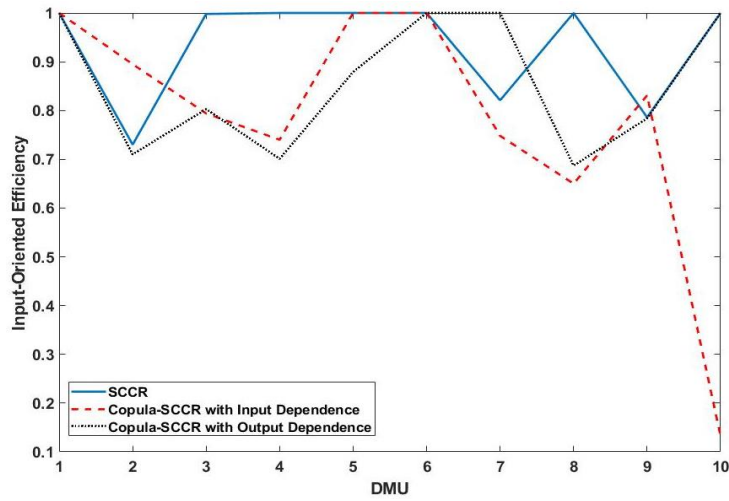


Figure 1. Performance of DMUs in input-oriented models.

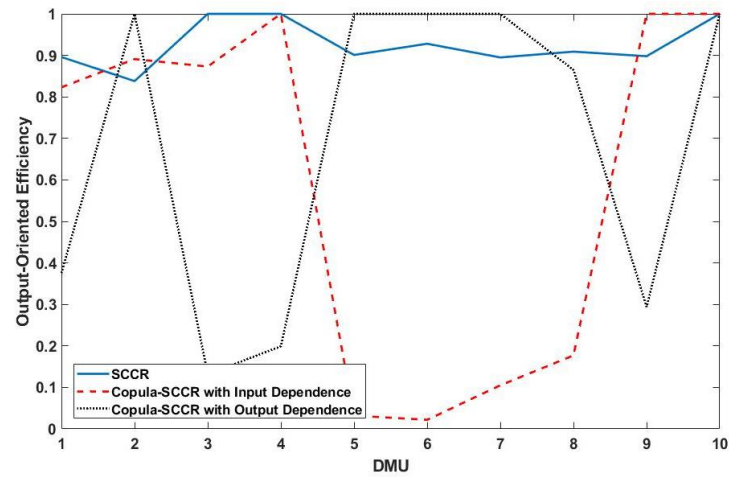


Figure 2. Performance of DMUs in output-oriented models.

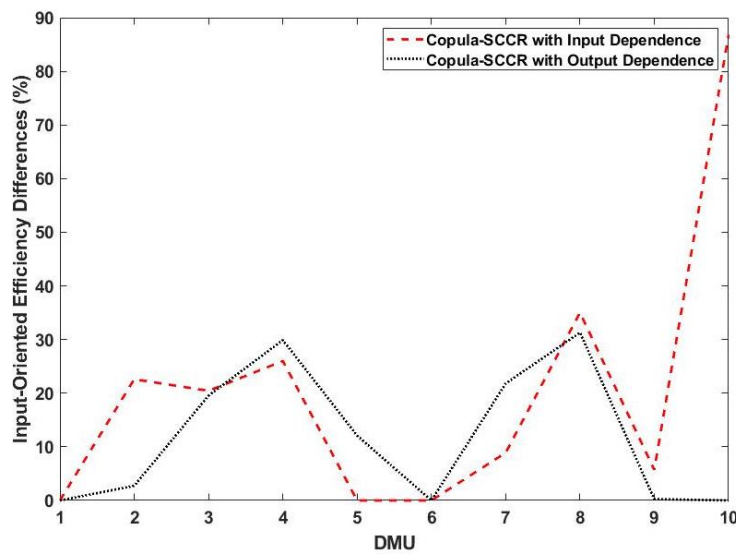
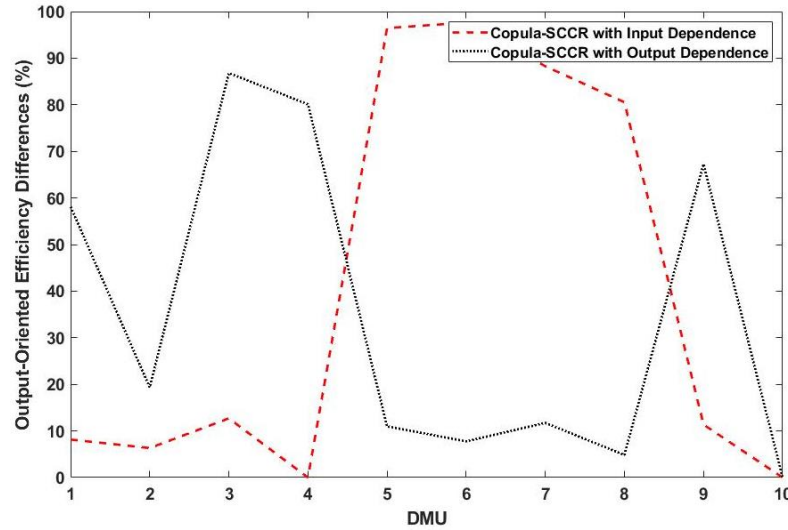


Figure 3. The difference between input-oriented Copula-SCCR models and input-oriented SCCR models



**Figure 4.** The difference between output-oriented Copula-SCCR models and output-oriented SCCR models

**Table 8.** Differences in mean values between the Copula-SCCR model and the SCCR model

Model	Mean Difference with SCCR Model (%)	
	Input-Oriented	Output-Oriented
Copula-SCCR with input dependence structure	20.546	40.140
Copula-SCCR with output dependence structure	11.764	34.696

## 6. Conclusion

When the input and output variables are not deterministic, but stochastic, DEA with stochastic data provides a powerful method for determining the efficiency of homogenous decision-making units. Probability distributions play a key role in analysing such data. There may be a dependence structure in stochastic variables. Failure to consider a dependence structure leads to incorrect results, and consequently to erroneous decisions in stochastic data envelopment analysis.

We assumed that the inputs and outputs are normally distributed in this study. Input- and output-oriented Copula-SCCR models were presented considering the dependence structure between the random variables. As the variables are normally distributed, we used the Gaussian copula to estimate the dependence structure. We evaluated the proposed models using the real data of 10 car companies. This study shows that, first, input-oriented and output-oriented Copula-SCCR models have different estimated efficiencies compared with the SCCR model. Second, in the output-oriented model, considering the dependence structure is far more important than considering the input-oriented model. Third, for both input- and output-oriented models, the Copula-SCCR model with an input dependence structure showed the most difference from the SCCR model. This is the very important finding that shows both in input- and output-oriented models, it is more important to account for the dependence structure between inputs than to consider the dependence structure between outputs. We recommended for future work to consider the dependence structure for other distributions, such as skew normal distribution, and also use these models in fuzzy DEA to get more accurate results.

**Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.



## References

1. Abdullah, D., & Erliana, C. I. (2021). Hesitant Fuzzy-Stochastic Data Envelopment Analysis (HF-SDEA) Model for Benchmarking. *JOIV: International Journal on Informatics Visualization*, 5(1), 94-98.
2. Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Management Science*, 30(9), 1078-1092.
3. Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), 429-444.
4. Cooper, W. W., Deng, H., Huang, Z., & Li, S. X. (2004). Chance constrained programming approaches to congestion in stochastic data envelopment analysis. *European Journal of Operational Research*, 155(2), 487-501.
5. Czado, C. (2019). Analyzing dependent data with vine copulas. *Lecture Notes in Statistics*, Springer.
6. Dibachi, H., Behzadi, M. H., & Izadikhah, M. (2015). Stochastic modified MAJ model for measuring the efficiency and ranking of DMUs. *Indian Journal of Science and Technology*, 8(8), 549-555.
7. El-Demerdash, B. E., El-Khodary, I. A., & Tharwat, A. A. (2013). Developing a stochastic input oriented data envelopment analysis (SIODEA) model. *International Journal of Advanced Computer Science and Applications*, 4(4), 40-45.
8. Gholami, O. (2021). A Novel Approach for Solving Fuzzy Stochastic Data envelopment analysis model in the presence of undesirable outputs. *Fuzzy Optimization and Modeling Journal*, 3(1), 22-36.
9. Hofert, M., Mächler, M., & McNeil, A. J. (2012). Likelihood inference for Archimedean copulas in high dimensions under known margins. *Journal of Multivariate Analysis*, 110, 133-150.
10. Hosseinzadeh Lotfi, F., Nematollahi, N., Behzadi, M. H., Mirbolouki, M., & Moghaddas, Z. (2012). Centralized resource allocation with stochastic data. *Journal of Computational and Applied Mathematics*, 236(7), 1783-1788.
11. Huang, Z., & Li, S. X. (2001). Stochastic DEA models with different types of input-output disturbances. *Journal of Productivity Analysis*, 15(2), 95-113.
12. Jradi, S., & Ruggiero, J. (2019). Stochastic data envelopment analysis: A quantile regression approach to estimate the production frontier. *European Journal of Operational Research*, 278(2), 385-393.
13. Kheirollahi, H., Matin, B. K., Mahboubi, M., & Alavijeh, M. M. (2015). Chance constrained input relaxation to congestion in stochastic DEA. An application to Iranian hospitals. *Global journal of health science*, 7(4), 151.
14. Khodabakhshi, M., & Asgharian, M. (2009). An input relaxation measure of efficiency in stochastic data envelopment analysis. *Applied Mathematical Modelling*, 33(4), 2010-2023.
15. Khodabakhshi, M., Asgharian, M., & Gregoriou, G. N. (2010). An input-oriented super-efficiency measure in stochastic data envelopment analysis: Evaluating chief executive officers of US public banks and thrifts. *Expert Systems with Applications*, 37(3), 2092-2097.
16. Land, K. C., Lovell, C. K., & Thore, S. (1993). Chance-constrained data envelopment analysis. *Managerial and Decision Economics*, 14(6), 541-554.
17. Li, F., Wu, L., Zhu, Q., Yu, Y., Kou, G., & Liao, Y. (2020). An eco-inefficiency dominance probability approach for Chinese banking operations based on data envelopment analysis. *Complexity*, 2020, Article ID 3780232.
18. Li, S. X. (1998). Stochastic models and variable returns to scales in data envelopment analysis. *European Journal of Operational Research*, 104(3), 532-548.
19. Ma, J. (2019). Discovering Association with Copula Entropy. *arXiv preprint arXiv:1907.12268*.
20. Miles, J. (2014). Wiley StatsRef: Statistics Reference Online. *R squared, adjusted R squared*, edited by N. Balakrishnan, T. Colton, B. Everitt, W. Piegorisch, F. Ruggeri, and JL Teugels. Chichester, UK: Wiley.
21. Morita, H., & Seiford, L. M. (1999). Characteristics on stochastic DEA efficiency: Reliability and probability being efficient. *Journal of the Operations Research Society of Japan*, 42(4), 389-404.
22. Nasser, S. H., Ebrahimnejad, A., & Gholami, O. (2016). Fuzzy stochastic input-oriented primal data envelopment analysis models with application to insurance industry. *International Journal of Applied Decision Sciences*, 9(3), 259-282.
23. Nasser, S. H., Ebrahimnejad, A., & Gholami, O. (2018). Fuzzy stochastic data envelopment analysis with undesirable outputs and its application to banking industry. *International Journal of Fuzzy Systems*, 20(2), 534-548.
24. Nazari, A., & Behzadi, M. H. (2017). Asymptotic distribution of the sum of skew-normal random variables: Application in data envelopment analysis. *Iranian Journal of Science and Technology, Transactions A: Science*, 41(1), 199-207.
25. Peykani, P., Farzipoor Saen, R., Seyed Esmaeili, F.S., & Gheidar-Kheljani, J. (2021). Window data envelopment analysis approach: a review and bibliometric analysis. *Expert Systems*, 38(7), e12721.
26. Peykani, P., Lotfi, F. H., Sadjadi, S. J., Ebrahimnejad, A., & Mohammadi, E. (2021). Fuzzy chance-constrained data envelopment analysis: a structured literature review, current trends, and future directions. *Fuzzy Optimization and Decision Making*, 1-65.
27. Peykani, P., Mohammadi, E., & Emrouznejad, A. (2021). An adjustable fuzzy chance-constrained network DEA approach with application to ranking investment firms. *Expert Systems with Applications*, 166, 113938.
28. Peykani, P., Mohammadi, E., Saen, R. F., Sadjadi, S. J., & Rostamy- Malkhalifeh, M. (2020). Data envelopment analysis and robust optimization: A review. *Expert Systems*, 37(4), e12534.
29. Peykani, P., Namakshenas, M., Arabjazi, N., Shirazi, F., & Kavand, N. (2021). Optimistic and pessimistic fuzzy data envelopment analysis: Empirical evidence from Tehran stock market. *Fuzzy Optimization and Modeling Journal*, 3(2), 12-21.
30. Santos Arteaga, F. J., Ebrahimnejad, A., & Zabihi, A. (2021). A new approach for solving intuitionistic fuzzy data envelopment analysis problems. *Fuzzy Optimization and Modeling Journal*, 2(2), 46-56.
31. Sengupta, J. K. (1987). Data envelopment analysis for efficiency measurement in the stochastic case. *Computers & Operations Research*, 14(2), 117-129.
32. Sengupta, J. K. (2000). Efficiency analysis by stochastic data envelopment analysis. *Applied Economics Letters*, 7(6), 379-383.

33. Sklar, L. (1995). *Physics and chance: Philosophical issues in the foundations of statistical mechanics*. Cambridge University Press.
34. Tavana, M., Shiraz, R. K., Hatami-Marbini, A., Agrell, P. J., & Paryab, K. (2012). Fuzzy stochastic data envelopment analysis with application to base realignment and closure (BRAC). *Expert Systems with Applications*, 39(15), 12247-12259.
35. Tavassoli, M., Saen, R. F., & Zanjirani, D. M. (2020). Assessing sustainability of suppliers: A novel stochastic-fuzzy DEA model. *Sustainable Production and Consumption*, 21, 78-91.
36. Thanassoulis, E., & Dunstan, P. (1994). Guiding schools to improved performance using data envelopment analysis: An illustration with data from a local education authority. *Journal of the Operational Research Society*, 45(11), 1247-1262.
37. Vaninsky, A. (2013). Stochastic DEA with a perfect object and its application to analysis of environmental efficiency.
38. Wanke, P., Barros, C. P., & Emrouznejad, A. (2018). A comparison between stochastic DEA and fuzzy DEA approaches: revisiting efficiency in Angolan banks. *RAIRO-Operations Research*, 52(1), 285-303.
39. Wei, G., Chen, J., & Wang, J. (2014). Stochastic efficiency analysis with a reliability consideration. *Omega*, 48, 1-9.



Balak, S., Behzadi, M. H., & Nazari, A. (2021). Comparison of the Dependence Structures of Stochastic Copula-DEA Model. *Fuzzy Optimization and Modelling Journal*, 2(4), 20-37.

[10.30495/fomj.2021.1943130.1041](https://doi.org/10.30495/fomj.2021.1943130.1041)

Received: 21 October 2021

Revised: 8 November 2021

Accepted: 9 November 2021



Licensee Fuzzy Optimization and Modelling Journal. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).