Theory of Approximation and Applications

Vol. 12, No.1, (2018), 93-115

Stochastic Multiplicative DEA for Estimating Most Productive Scale Size

Hossein Dibachi [a,*]

[a] *Department of Mathematics, Arak Branch, Islamic Azad University, Arak, Iran*

Received 01 July 2018; accepted 03 October 2018

Abstract

In this paper, stochastic multiplicative data envelopment analysis (MDEA) model under variable return to scale (VRS) technology in the presence of log-normal distribution is proposed for estimating most productive scale size (MPSS). Banker and Maindiratta introduced MPSS pattern in MDEA model. The MDEA model requires that the values for all inputs and outputs be known exactly. But this assumption is not always correct, because data in many practical situations cannot be precisely measured. One of the most important methods, when we're dealing with imprecise data is considering stochastic data. Moreover, for solving stochastic model, a deterministic equivalent is obtained and also stochastic $\alpha-$MPSS is defined for decision making units (DMUs). Finally, an example of the systems reliability is presented to demonstrate our proposed modeling idea and its efficiency.

Key words: Log-normal distribution; Multiplicative data envelopment analysis(MDEA); Stochastic MDEA; Stochastic $\alpha-$MPSS; Systems reliability.

---

* Corresponding author's e-mail: h-dibachi@iau-arak.ac.ir(H.Dibachi)

1 Introduction

Data envelopment analysis (DEA) involves an alternative principle for extracting information about a population of observations called decision making units (DMUs) with similar quantitative characteristics. This is reflected by the assumption that each DMU uses the same set of inputs to produce the same set of outputs, but the inputs are consumed and outputs are produced in varying amounts.

The first DEA model (CCR model) that successfully optimized each individual observation, DMU, with the objective of calculating a discrete piecewise frontier was proposed by Charnes et al. (1978) and extended by Banker et al. (1984). One class of models introduced in DEA is called multiplicative data envelopment analysis (MDEA) model, in which, as shown by Banker and Maindiratta (1986), the piecewise linear frontiers usually employed in DEA are replaced by a frontier that is piecewise Cobb-Douglas. Banker and Maindiratta (1986), introduced a model to identify the most productive scale size (MPSS) pattern, and Banker et al. (2004) presented a two-stage method for the identification of returns to scale in MDEA model. In the BCC model the convexity postulate permits increasing, constant or decreasing returns to scale in different regions of the production function.

However, this also requires the marginal products (see, Menger (1954), for a comparison of returns to scale and rate of change of marginal product) to be nonincreasing. This restriction in the BCC approach may not be appropriate for production technologies where the production function is nonconcave in some regions and the production possibility set is not convex. To allow for such situations, Banker and Maindiratta (1986), replace the ordinary convexity postulate of BCC by "geometric" convexity to interpolate between observed production possibilities. This implies that the piecewise linear frontiers, usually employed in DEA, are replaced by a frontier that is piecewise log-linear (Zarepisheh et al., 2009; Mehdiloozad et al., 2014). Since the introduction of DEA, there has been an impressive growth both in its theoretical developments and applications (Hollingsworth et al., 1999; Cook and Seiford, 2009; Cooper et al., 2006; Cao and Yang, 2011).

Reader can also refer to Xing et al. (2013) where some applications of DEA in service industries are mentioned. The economical concept of returns to scale has also been widely studied within the DEA framework. If in an empirical application there are a priori reasons to believe that marginal products are increasing in some regions, then the log-linear model is the appropriate DEA model for the analysis. Banker et al. (1981) describe a procedure for piecewise log-linear estimation of the efficient production surface. Then, Charnes et al. (1982) employed this log-linear envelopment principle in Banker et al. (1981) to suggest a multiplicative efficiency measure. For more details about the multiplicative models and applications, see e.g., Chang and Guh (1994), Charnes, Cooper, Seiford, and Stutz (1983), Seiford and Zhu (1998), Sueyoshi and Chang (1989), Zarepisheh et al. (2010) and Davoodi et al. (2015). Classic DEA models do not allow stochastic variations in input-output data, such as measurement errors and data entry errors. In traditional form of DEA models, the data of inputs and outputs of the different DMUs are assumed to be measured with precision. On the other hand, this is not always possible. For removing this weakness in the classic DEA models, some authors proposed stochastic input and output variations into the DEA. The stochastic data envelopment analysis (SDEA) approach was developed by considering the value of inputs and outputs as random variables. Banker (1993), for example, incorporated the statistical elements into the DEA and developed a nonparametric approach with maximum likelihood methods to effect inferences in the presence of statistical noise. Olesen and Petersen (1995) developed a chance-constrained DEA model which used the piecewise linear envelopments of confidence regions for use with stochastic multiple inputs and multiple outputs.

Cooper et al. (1998) developed a“joint chance-constrained ” DEA model to naturally generalize “ Pareto-Koopman’s Efficiency” to stochastic situations. Huang and Li (1996) utilized this joint chance-constrained concept to discuss general dominance structures in the stochastic situations. Cooper et al. (2002, 2003) have introduced the chance-constrained models to deal with the technical inefficiencies and congestion in the stochastic situation. Land et al. (1993) presented an alternative chance-constrained formulation of DEA, starting out from the multiplicative model and assuming that the joint probability distribution of all outputs is log-normal.

Khodabakhshi (2009) input-output oriented model which was first introduced by Jahanshahloo and Khodabakhshi (2003), developed in stochastic data envelopment analysis to identify MPSS units and assuming that the all input and output components are jointly normally distributed.

Lee (2015) proposed a multi-objective mathematical program with DEA constraints to set an efficient target which shows a trade-off between the MPSS benchmark and a potential demand fulfillment benchmark. When dealing with failure and repair mechanisms in general, the most suitable and applied distribution is the log-normal distribution. Therefore, in this paper, we propose the stochastic input-output oriented MDEA model under VRS technology for estimating stochastic MPSS pattern of systems. We consider these systems as DMUs with the inputs and outputs having log-normal distributions where inputs and outputs are stochastic repair times and stochastic failure times, respectively. This paper is structured as follows:

Some basic concepts in statistics, input-output BCC model and deterministic MDEA model will be introduced in the next section. Section 3 addresses the proposed method for estimating the stochastic MPSS in input-output stochastic MDEA model. A brief discuss about the proposed models and an numerical example in systems reliability are given in section 4. Conclusions will appear in section 5.

2 Preliminaries

In this section, we recall some basic concepts and results which will be used through the paper.

2.1 Log-normal Distribution

A random variable which is log-normally distributed takes only positive real values. In probability theory, a log-normal distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. The log-normal distribution is important in the description of natural phenomena. Some of these applications are as follows:

- In quantitative economics and finance, the log-normal distribution is ubiquitous and it arises, among other things, in connection with geometric Brownian motion, the standard model for the price dynamics of securities in mathematical finance.
- In finance, in particular the Black-Scholes model, changes in the logarithm of exchange rates, price indices, and stock market indices are assumed normal.
- A main area of application for the log-normal distribution is lifetime research and reliability theory.

Definition 2.1 *A random variable X is said to have the log-normal distribution if its probability density function is given as follows:*

$$f_X(x) = \begin{cases} \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(lnx-\mu)^2} & ; x > 0 \\ 0 & ; o.w \end{cases} \quad (2.1)$$

We will use the notation $X \sim LN(\mu, \sigma^2)$ to denote the random variable X having the log-normal distribution with parameters $\sigma > 0$ and $\mu \in \mathbf{R}$ where $\mu = E(LnX)$ and $\sigma^2 = Var(LnX)$.

Remark 2.1 *If $X \sim LN(\mu, \sigma^2)$, then $Y = LnX$ having the normal distribution with scale parameter $\sigma > 0$ and location parameter $\mu \in \mathbf{R}$ where is denoted by notation $Y \sim N(\mu, \sigma^2)$. Thus, probability density function of Y is given as follows:*

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y-\mu)^2} \; ; y \in \mathbf{R} \quad (2.2)$$

The corresponding cumulative distribution function has the following form:

$$F_Y(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} dt \quad (2.3)$$

Note that if $Y \sim N(0, 1)$ then $f_Y(y)$ is called standard normal distribution and $F_Y(y)$ is denoted by $\Phi(y)$ and Φ^{-1}, its inverse, is the so-called *fractile function. For example,* $\Phi^{-1}(0.1) = -1.28$*,* $\Phi^{-1}(0.33) = -0.44$*,* $\Phi^{-1}(0.5) = 0$*,* $\Phi^{-1}(0.67) = 0.44$*, and* $\Phi^{-1}(0.9) = 1.28$*.*

2.2 System Reliability

A system contains one or several subsystems of components, henceforth called items, interconnected so that the system is able to perform of number of required functions. The reliability of the system denotes the relationship between the systems required performance and its achieved performance. The probabilistic approach of the system's reliability deals with the uncertainty of this relation. To prevent system failures, e.g. failures that prevents the system from performing any of its supposed functions, the potential failures should be identified. To describe an item's characteristics in terms of reliability there are several functions that can be used. The failure rate function, $z(t)$ describes the components tendency to fail, failures per time unit, for $t \geq 0$. However, the instantaneous failure rate at the time t_0 for functional items rate is called $\gamma = z(t_0)$, the corresponding instantaneous repair rate for faulted items is called μ. In order to comprehend an item's stochastic behaviour concerning its uptime, functional, and downtime, faulted, the item's probabilistic behaviour can be represented using a distribution function (see, Stapelberg, 2009). In reliability analysis, failure time and repair time a system is often distributed log-normally.

2.3 Input-Output Oriented BCC Model

One of the basic DEA model for evaluating DMUs is the BCC model where introduced by Banker et al. (1984). They omitted the ray unboundedness postulate from the CCR postulates and deduced the following production possibility set:

$$T_{BCC} = \{(\mathbf{x},\mathbf{y}) \mid \ \mathbf{x} \geq \sum_{j=1}^{n} \lambda_j \mathbf{x}_j \ \& \mathbf{y} \leq \sum_{j=1}^{n} \lambda_j \mathbf{y}_j \& \ \sum_{j=1}^{n} \lambda_j = 1 \ \& \ \lambda_j \geq 0\} \tag{2.4}$$

Where $\mathbf{x}_j = (x_{1j}, x_{2j}, \ldots, x_{mj}) \in \mathbb{R}^m_{\geq 0}$ and $\mathbf{y}_j = (y_{1j}, y_{2j}, \ldots, y_{sj}) \in \mathbb{R}^s_{\geq 0}$ are the input and output vectors of DMU_j, respectively. Banker (1984), introduced the following model to identify the MPSS pattern for an efficient DMU_o in the input-output oriented BCC model

$$
\begin{aligned}
&\tfrac{\phi_o^*}{\theta_o^*} = Maximize \;\; \tfrac{\phi_o}{\theta_o} \\
&s.t. \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geq \phi_o y_{ro}, r = 1, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta_o x_{io}, i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, j = 1, \ldots, n
\end{aligned}
\tag{2.5}
$$

Using the above model Cooper et. al (1996) provided a theorem which defines MPSS as follows:

Definition 2.2 *DMU_o is said to be MPSS if and only if the following two conditions are both satisfied for model (2.5):*

i) $\frac{\phi_o^*}{\theta_o^*} = 1$
ii) *All slack variables are zero in the alternative optimal solution.*

Definition 2.3 *(Banker's Definition):$(X_o, Y_o) \in T$ is most productive scale size (MPSS) if and only if for every $(\theta_o X_o, \phi_o Y_o) \in T$ we have $\theta_o \geq \phi_o$.*

Note that the Cooper et al.'s definition of MPSS is stronger from Banker's definition, because of considering slacks in alternative optimal solutions, in other words they defined strong or Pareto-efficient MPSS.

Multiplicative data envelopment analysis (MDEA) model was first introduced by Charnes et al. (1982). Suppose that there are n DMUs, where each DMU_j (j=1,..., n) uses m different inputs, $x_{ij} > 0$ (i=1,..., m), to produce s different outputs, $y_{rj} > 0$ (r=1,..., s) and suppose also that the data set is deterministic. Therefore for each DMU_j, let $\mathbf{x}_j = (x_{1j}, x_{2j}, \ldots, x_{mj})$ and $\mathbf{y}_j = (y_{1j}, y_{2j}, \ldots, y_{sj})$ are the input and output vectors of DMU_j, respectively. A production technology transforming an input vector $\mathbf{x}_j \in \mathbb{R}^m_{>0}$ into an output vector $\mathbf{y}_j \in \mathbb{R}^s_{>0}$ can be characterized by the technology set T_M which is defined as follows:

$$T_M = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{m+s}_{>0} \mid \ \mathbf{x} \textit{ can be produced by } \mathbf{y}\} \tag{2.6}$$

We assume that T_M satisfies the following four postulates:

i) T_M is a closed set.
ii) For each $\mathbf{x} \in \mathbb{R}^m_{>0}$ the set $B(\mathbf{x}) = \{(\mathbf{u}, \mathbf{y}) \in T_M \mid \ \mathbf{u} \leq \mathbf{x}\}$ is bounded.
iii) T_M satisfies free disposability for all the inputs and outputs, i.e., if $(\mathbf{x}, \mathbf{y}) \in T_M$ and $(\mathbf{x}, -\mathbf{y}) \leq (\mathbf{x}', -\mathbf{y}')$ then $(\mathbf{x}', \mathbf{y}') \in T_M$.
iv) T_M has the geometric convexity, i.e., if $(\mathbf{x_1}, \mathbf{y_1}) \in T_M$ and $(\mathbf{x_2}, \mathbf{y_2}) \in T_M$ then $(\mathbf{x_1}^{\lambda}\mathbf{x_2}^{1-\lambda}, \mathbf{y_1}^{\lambda}\mathbf{y_2}^{1-\lambda}) \in T_M$ for all $\lambda \in [0, 1]$

To further clarify the Postulate (iv), we define the log form of T_M as

$$lnT_M = \{(ln\mathbf{x}, ln\mathbf{y}) \mid \ (\mathbf{x}, \mathbf{y}) \in T_M\}, \tag{2.7}$$

in which $ln\mathbf{x} = (lnx_1, lnx_2, \cdots, lnx_m)$ and $ln\mathbf{y} = (lny_1, lny_2, \cdots, lny_s)$. By using of the strict monotonicity property of the natural logarithm function, there is a one-to-one correspondence between T_M and lnT_M. Therefore, the geometric convexity of T_M is equivalent to the ordinary convexity of lnT_M. This establishes that T_M is geometric convex if and only if for all $(ln\mathbf{x_1}, ln\mathbf{y_1}), (ln\mathbf{x_2}, ln\mathbf{y_2}) \in lnT_M$, and all $\lambda \in [0, 1]$ the following condition is satisfied:

$$(\lambda ln\mathbf{x_1} + (1-\lambda)ln\mathbf{x_2}, \lambda ln\mathbf{y_1} + (1-\lambda)ln\mathbf{y_2}) \in lnT_M. \tag{2.8}$$

Assuming the inputs and outputs to be strictly positive, we define the piece-wise log-linear technology T_1 that is constructed from the observed DMUs under Postulates (i)–(iv). Since assuming geometric convexity for T_M is tantamount to assuming convexity for lnT_M, lnT_M will be piece-wise linear provided T_M is geometric convex, and as a result, T_1 is called piece-wise log-linear. Banker and Maindiratta (1986) replaced the ordinary convexity postulate of BCC by "geometric" convexity, and introduced the following production possibility set (PPS):

$$T_1 = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{m+s}_{>0} \mid \mathbf{x} \geq \prod_{j=1}^{n} \mathbf{x}_j^{\lambda_j} \ \& \ \mathbf{y} \leq \prod_{j=1}^{n} \mathbf{y}_j^{\lambda_j} \ \& \ \sum_{j=1}^{n} \lambda_j = 1 \ \& \ \lambda_j \geq 0\} \tag{2.9}$$

and

$$lnT_1 = \{(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{m+s}_{>0} \mid ln\mathbf{x} \geq \sum_{j=1}^{n} \lambda_j ln\mathbf{x}_j \ \& \ ln\mathbf{y} \leq \sum_{j=1}^{n} \lambda_j ln\mathbf{y}_j \ \& \ \sum_{j=1}^{n} \lambda_j = 1 \ \& \ \lambda_j \geq 0\} \tag{2.10}$$

Note that T_1 is free from this restriction, and allows for increasing, constant and decreasing marginal products. For details of the empirical technological structures of T_1 and lnT_1, see Mehdiloozad et al. (2014). Therefore, output-oriented MDEA model under VRS technology for evaluating DMU_o is given by the following model:

$$\begin{aligned}
&\phi_o^* = Maximize \ \ \phi_o \\
&s.t. \\
&\prod_{j=1}^{n} y_{rj}^{\lambda_j} \geq \phi_o y_{ro}, r = 1, \ldots, s, \\
&\prod_{j=1}^{n} x_{ij}^{\lambda_j} \leq x_{io}, i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, j = 1, \ldots, n
\end{aligned} \tag{2.11}$$

With the above assumptions to convert these inequalities to equations we setting $\phi_o y_{ro} = e^{-s_r^+} \prod_{j=1}^{n} y_{rj}^{\lambda_j}$ and $x_{io} = e^{s_i^-} \prod_{j=1}^{n} x_{ij}^{\lambda_j}$ and also replace the objective in (2.11) with $\phi_o exp(\varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+))$. Now, by taking the natural logarithm of both sides, in the first and second constraint in model (2.11) non-Archimedean model is stated as follows:

$$
\begin{aligned}
&\tilde{\phi}_o^* = \ Maximize \ \ \tilde{\phi}_o + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+) \\
&s.t. \\
&\sum_{j=1}^{n} \lambda_j \tilde{y}_{rj} - s_r^+ = \tilde{\phi}_o + \tilde{y}_{ro}, r = 1, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j \tilde{x}_{ij} + s_i^- = \tilde{x}_{io}, i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, s_r^+ \geq 0, s_i^- \geq 0, j = 1, \ldots, n
\end{aligned}
\tag{2.12}
$$

in which " $\sim$ " denotes "natural logarithm" and also $s_i^- \geq 0$ and $s_r^+ \geq 0$ represent slacks. The optimal value $\phi_o^* = e^{\tilde{\phi}_o^*}$ obtained from the linear programming formulation in (2.12).

Definition 2.4 *DMU_o is said to be efficient if and only if the following two conditions are both satisfied for model (2.12):*

i) $e^{\tilde{\phi}_o^*} = 1.$
ii) *All slack variables are zero in the alternative optimal solution.*

Similarly, input-oriented MDEA model under VRS technology is stated.

Banker and Maindiratta (1986), introduced the following model to identify the MPSS pattern for an efficient DMU_o

$$
\begin{aligned}
&\frac{\phi_o^*}{\theta_o^*} = Maximize \ \ \frac{\phi_o}{\theta_o} \\
&s.t. \\
&\prod_{j=1}^{n} y_{rj}^{\lambda_j} \geq \phi_o y_{ro}, r = 1, \ldots, s, \\
&\prod_{j=1}^{n} x_{ij}^{\lambda_j} \leq \theta_o x_{io}, i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, j = 1, \ldots, n
\end{aligned}
\tag{2.13}
$$

They showed that $(\theta_o^* \mathbf{x}_o, \phi_o^* \mathbf{y}_o)$ is a MPSS pattern for DMU_o. Similar to the steps for determining model (2.12) we setting $\phi_o y_{ro} = e^{-s_r^+} \prod_{j=1}^{n} y_{rj}^{\lambda_j}$ and $\theta_o x_{io} = e^{s_i^-} \prod_{j=1}^{n} x_{ij}^{\lambda_j}$ and also replace the objective in (2.13) with $\frac{\phi_o}{\theta_o} exp(\varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+))$. Now, by taking the natural logarithm of both sides, in the first and second constraint in model (2.13) non-Archimedean model is stated as follows:

$$
\begin{aligned}
&\tilde{\phi}_o^* - \tilde{\theta}_o^* = Maximize \ \ \tilde{\phi}_o - \tilde{\theta}_o + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+) \\
&s.t. \\
&\sum_{j=1}^{n} \lambda_j \tilde{y}_{rj} - s_r^+ = \tilde{\phi}_o + \tilde{y}_{ro}, r = 1, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j \tilde{x}_{ij} + s_i^- = \tilde{\theta}_o + \tilde{x}_{io}, i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, s_r^+ \geq 0, s_i^- \geq 0, j = 1, \ldots, n
\end{aligned}
\tag{2.14}
$$

As in axiomatic approach of Banker et al. (1984), we can axiomatically derive the formulation in model (2.14) by adding the "Ray Extension" (i.e. if $(\mathbf{x}, \mathbf{y}) \in T_M$ and $t > 0$ then $(t\mathbf{x}, t\mathbf{y}) \in T_M$)) Postulate to the set of Postulate (i)–(iv) we can then also obtain the optimal value $\frac{\phi_o^*}{\theta_o^*} = e^{\tilde{\phi}_o^* - \tilde{\theta}_o^*}$ from the linear programming formulation in (2.14).

Definition 2.5 *DMU_o is said to be MPSS if and only if the following two conditions are both satisfied for model (2.14):*

i) $e^{\tilde{\phi}_o^* - \tilde{\theta}_o^*} = 1$
ii) *All slack variables are zero in the alternative optimal solution.*

For solving model (2.14) at first without any attention to slacks we obtain $Maximum(\tilde{\phi}_o - \tilde{\theta}_o)$, and then in the second stage we maximize slacks by fixing $\tilde{\phi}_o^*$ and $\tilde{\theta}_o^*$ values instead of $\tilde{\phi}_o$ and $\tilde{\theta}_o$ under their corresponding constraints. Note that in this approach there is no need to determine any value for ε.

3 Stochastic Most Productive Scale Size Pattern in Multiplicative Model

In this section, input-output oriented multiplicative model introduced by Banker and Maindiratta (1986) is developed into stochastic multiplicative DEA to identify most productive scale size units. Throughout this article, random variables are denoted by capital letters. For each $DMU_j, (j = 1, \ldots, n)$, let $\mathbf{X}_j = (X_{1j}, X_{2j}, \ldots, X_{mj}) \in \mathbb{R}^m_{>0}$ and $\mathbf{Y}_j = (Y_{1j}, Y_{2j}, \ldots, Y_{sj}) \in \mathbb{R}^s_{>0}$ are the input and output random vectors of DMU_j, respectively. Suppose that all input and output components are jointly Log-normally distributed, i.e., $X_{ij} \sim LN(\mu_{ij}, \sigma^2_{ij}), (i = 1, \ldots, m)$ and $Y_{rj} \sim LN(\gamma_{rj}, \tau^2_{rj}), (r = 1, \ldots, s)$. By Remark (2.1) $LnX_{ij} = \tilde{X}_{ij} \sim N(\mu_{ij}, \sigma^2_{ij})$ and $LnY_{rj} = \tilde{Y}_{rj} \sim N(\gamma_{rj}, \tau^2_{rj})$. Now, by using model (2.13) the stochastic input-output oriented multiplicative model to identify the stochastic MPSS pattern for an efficient DMU_o is proposed as follows:

$$\begin{aligned}
&\tfrac{\phi^*_o(\alpha)}{\theta^*_o(\alpha)} = Maximize \ \ \tfrac{\phi_o}{\theta_o} \\
&s.t. \\
&P(\prod_{j=1}^{n} Y_{rj}^{\lambda_j} \geq \phi_o Y_{ro}) \geq 1 - \alpha, r = 1, \ldots, s, \\
&P(\prod_{j=1}^{n} X_{ij}^{\lambda_j} \leq \theta_o X_{io}) \geq 1 - \alpha i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, j = 1, \ldots, n
\end{aligned} \tag{3.1}$$

where α is a predetermined number between 0 and 1 which specifies the significance level and P means “Probability Measure”. Since a solution with $\theta_o = \phi_o = 1, \lambda_o = 1, \lambda_j = 0$ $(j \neq o)$, always exists, the optimal value of objective function is greater than or equal to one. The corresponding stochastic version of the model (3.1), including slack variables, is stated as follows:

$$
\begin{aligned}
&\tilde{\phi}_o^*(\alpha) - \tilde{\theta}_o^*(\alpha) = Maximize \ \ \tilde{\phi}_o - \tilde{\theta}_o + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+) \\
&s.t. \\
&P(\sum_{j=1}^{n} \lambda_j \tilde{Y}_{rj} - s_r^+ - \tilde{Y}_{ro} \geq \tilde{\phi}_o) = 1 - \alpha, r = 1, \ldots, s, \\
&P(\sum_{j=1}^{n} \lambda_j \tilde{X}_{ij} + s_i^- - \tilde{X}_{io} \leq \tilde{\theta}_o) = 1 - \alpha, i = 1, \ldots, m, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, s_r^+ \geq 0, s_i^- \geq 0, j = 1, \ldots, n
\end{aligned}
\tag{3.2}
$$

3.1 Deterministic Equivalent

In this section, we utilize the log-normality assumption to introduce a deterministic equivalent to the model (3.2). If $X_{ij} \sim LN(\mu_{ij}, \sigma_{ij}^2)$ and $Y_{rj} \sim LN(\gamma_{rj}, \tau_{rj}^2)$ then $\tilde{X}_{ij} \sim N(\mu_{ij}, \sigma_{ij}^2)$ and $\tilde{Y}_{rj} \sim N(\gamma_{rj}, \tau_{rj}^2)$. Therefore, for all $r \in \{1, 2, \ldots, s\}$, $o \in \{1, 2, \ldots, n\}$, and $i \in \{1, 2, \ldots, m\}$ we have:

$$
\begin{aligned}
\sigma_i^2(\lambda) &= Var(\sum_{j=1}^{n} \lambda_j \tilde{X}_{ij} - \tilde{X}_{io} + s_i^- - \tilde{\theta}_o) \\
&= \sum_{k=1, k\neq o}^{n} \sum_{j=1, j\neq o}^{n} \lambda_j \lambda_k Cov(\tilde{X}_{ik}, \tilde{X}_{ij}) \\
&\quad + 2(\lambda_o - 1) \sum_{j=1, j\neq o}^{n} \lambda_j Cov(\tilde{X}_{ij}, \tilde{X}_{io}) + (\lambda_o - 1)^2 \sigma_{io}^2
\end{aligned}
\tag{3.3}
$$

Similarly,

$$\tau_r^2(\lambda) = Var(\sum_{j=1}^{n} \lambda_j \tilde{Y}_{rj} - \tilde{Y}_{ro} - s_r^+ - \tilde{\phi}_o)$$

$$\begin{aligned} &= \sum_{k=1,k\neq o}^{n} \sum_{j=1,j\neq o}^{n} \lambda_j \lambda_k Cov(\tilde{Y}_{rk}, \tilde{Y}_{rj}) \\ &\quad +2(\lambda_o - 1) \sum_{j=1,j\neq o}^{n} \lambda_j Cov(\tilde{Y}_{rj}, \tilde{Y}_{ro}) + (\lambda_o - 1)^2 \tau_{ro}^2 \end{aligned} \tag{3.4}$$

Using this results, can obtain the deterministic equivalent of model (3.2).

Theorem 3.1 *Deterministic equivalent of model (3.2) is as follows:*

$$\begin{aligned}
&\tilde{\phi}_o^*(\alpha) - \tilde{\theta}_o^*(\alpha) = Max \quad \tilde{\phi}_o - \tilde{\theta}_o + \varepsilon(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+) \\
&s.t. \\
&\sum_{j=1}^{n} \lambda_j \gamma_{rj} - \gamma_{ro} - s_r^+ + \tau_r(\lambda)\Phi^{-1}(\alpha) = \tilde{\phi}_o, r = 1, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j \mu_{ij} - \mu_{io} + s_i^- - \sigma_i(\lambda)\Phi^{-1}(\alpha) = \tilde{\theta}_o, i = 1, \ldots, m, \\
&\sigma_i^2(\lambda) = \sum_{k=1,k\neq o}^{n} \sum_{j=1,j\neq o}^{n} \lambda_j \lambda_k Cov(\tilde{X}_{ik}, \tilde{X}_{ij}) \\
&+2(\lambda_o - 1) \sum_{j=1,j\neq o}^{n} \lambda_j Cov(\tilde{X}_{ij}, \tilde{X}_{io}) + (\lambda_o - 1)^2 \sigma_{io}^2, i = 1, \ldots, m, \\
&\tau_r^2(\lambda) = \sum_{k=1,k\neq o}^{n} \sum_{j=1,j\neq o}^{n} \lambda_j \lambda_k Cov(\tilde{Y}_{rk}, \tilde{Y}_{rj}) \\
&+2(\lambda_o - 1) \sum_{j=1,j\neq o}^{n} \lambda_j Cov(\tilde{Y}_{rj}, \tilde{Y}_{ro}) + (\lambda_o - 1)^2 \tau_{ro}^2, r = 1, \ldots, s, \\
&\sum_{j=1}^{n} \lambda_j = 1, \\
&\lambda_j \geq 0, s_r^+ \geq 0, s_i^- \geq 0, j = 1, \ldots, n
\end{aligned} \tag{3.5}$$

Proof. From the first constraint in model (3.2) and Equation (3.4) we have:

$$P(\sum_{j=1}^{n}\lambda_j\tilde{Y}_{rj}-s_r^{+}-\tilde{Y}_{ro}-\tilde{\phi}_o\geq 0)=1-\alpha \iff$$

$$P(\frac{\sum_{j=1}^{n}\lambda_j\tilde{Y}_{rj}-\tilde{Y}_{ro}-s_r^{+}-\tilde{\phi}_o-(\sum_{j=1}^{n}\lambda_j\gamma_{rj}-\gamma_{ro}-s_r^{+}-\tilde{\phi}_o)}{\tau_r(\lambda)}$$

$$\geq\frac{-\sum_{j=1}^{n}\lambda_j\gamma_{rj}+\gamma_{ro}+s_r^{+}+\tilde{\phi}_o}{\tau_r(\lambda)})=1-\alpha \iff$$

$$P(Z\geq\frac{-\sum_{j=1}^{n}\lambda_j\gamma_{rj}+\gamma_{ro}+s_r^{+}+\tilde{\phi}_o}{\tau_r(\lambda)})=1-\alpha \iff$$

$$\Phi(\frac{-\sum_{j=1}^{n}\lambda_j\gamma_{rj}+\gamma_{ro}+s_r^{+}+\tilde{\phi}_o}{\tau_r(\lambda)})=\alpha \iff$$

$$\Phi^{-1}(\alpha)=\frac{-\sum_{j=1}^{n}\lambda_j\gamma_{rj}+\gamma_{ro}+s_r^{+}+\tilde{\phi}_o}{\tau_r(\lambda)} \iff \tag{3.6}$$

$$\sum_{j=1}^{n}\lambda_j\gamma_{rj}-s_r^{+}-\gamma_{ro}+\tau_r(\lambda)\Phi^{-1}(\alpha)=\tilde{\phi}_o$$

Similarly, from the second constraint in model (3.2) and Equation (3.3) we have:

$$P(\sum_{j=1}^{n}\lambda_j\tilde{X}_{ij}-\tilde{X}_{io}+s_i^- -\tilde{\theta}_o\le 0)=1-\alpha \qquad\Longleftrightarrow$$

$$P(\frac{\sum_{j=1}^{n}\lambda_j\tilde{X}_{ij}-\tilde{X}_{io}+s_i^- -\tilde{\theta}_o-(\sum_{j=1}^{n}\lambda_j\mu_{ij}-\mu_{io}+s_i^- -\tilde{\theta}_o)}{\sigma_i(\lambda)}\le$$

$$\frac{-\sum_{j=1}^{n}\lambda_j\mu_{ij}+\mu_{io}-s_i^- +\tilde{\theta}_o}{\sigma_i(\lambda)})=1-\alpha \qquad\Longleftrightarrow$$

$$P(Z\le\frac{-\sum_{j=1}^{n}\lambda_j\mu_{ij}+\mu_{io}-s_i^- +\tilde{\theta}_o}{\sigma_i(\lambda)})=1-\alpha \qquad\Longleftrightarrow$$

$$\Phi(\frac{\sum_{j=1}^{n}\lambda_j\mu_{ij}-\mu_{io}+s_i^- -\tilde{\theta}_o}{\sigma_i(\lambda)})=\alpha \qquad\Longleftrightarrow$$

Therefore,

$$\Phi^{-1}(\alpha)=\frac{\sum_{j=1}^{n}\lambda_j\mu_{ij}-\mu_{io}+s_i^- -\tilde{\theta}_o}{\sigma_i(\lambda)} \qquad\Longleftrightarrow \qquad (3.7)$$

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

Thus, by (3.6) and (3.7), the deterministic model is completely specified. □

Stochastic $\alpha-MPSS$ by solving model (3.5) can be defined as follows:

Definition 3.1 *DMU_o is said to be stochastic $\alpha-MPSS$ if and only if the following two conditions are both satisfied for model (3.5):*

i) $e^{\tilde{\phi}_o^*(\alpha)-\tilde{\theta}_o^*(\alpha)}=1$
ii) *All slack variables are zero in the alternative optimal solution.*

In Definition (3.1), if for an optimal solution, $e^{\tilde{\phi}_o^*(\alpha)-\tilde{\theta}_o^*(\alpha)}\neq 1$, or some of slacks are non zero, then DMU_o is not stochastic $\alpha-$MPSS. DMUs which are only satisfied in condition (i) are called weakly stochastic $\alpha-$MPSS.

Remark 3.1 *If $\alpha = 0.5$, then $\Phi^{-1}(0.5) = 0$. Therefore, the MPSS classification of DMU_o in input-output orientation MDEA model (2.13) is the same as in stochastic input-output orientation MDEA model (3.1) in which the mean values of inputs and outputs are used.*

4 Numerical Example of System Reliability

We apply the proposed stochastic MPSS pattern in input-output stochastic MDEA methodology for estimating the stochastic $\alpha-$MPSS of 12 systems. We consider this systems as DMUs, and denote them by DMU_j, ($j = 1, 2, \ldots, 12$). Every DMU_j is composed of 2 components (or items) which have the random length of time until failure. Suppose that Y_{rj} , $r = 1, 2$ are the random failure time of component r of DMU_j where have the log-normal distribution with parameters τ^2_{rj} and γ_{rj} which are denoted with $Y_{rj} \sim LN(\gamma_{rj}, \tau^2_{rj})$. When a component fails it undergoes repair. Suppose that X_{ij} , $i = 1, 2$ are the random repair time of component i of DMU_j where have the log-normal distribution with parameters σ^2_{ij} and μ_{ij} which are denoted with $X_{ij} \sim LN(\mu_{ij}, \sigma^2_{ij})$. Thus, by solving model (3.5) can be obtain the stochastic $\alpha-$MPSS of systems. The labels of inputs and outputs are as Table 1. The data set for this example

Table 1
The labels of inputs and outputs.

Input1:	The random repair time of the first component of system
Input2:	The random repair time of the second component of system
Output1:	The random length of time until failure of the first component of system
Output2:	The random length of time until failure of the second component of system

is shown in Table 2. We run model (3.5) by means of GAMS software for all $\alpha \in \{0.1, 0.33, 0.5, 0.67, 0.9\}$ and the results are shown in Table 3. In Table 3, DMU_1 and DMU_2 have optimal solutions $e^{\tilde{\phi}^*_o(\alpha)-\tilde{\theta}^*_o(\alpha)} = 1$, $s_1^{-*} = 0$, $s_2^{-*} = 0$, $s_1^{+*} = 0$, and $s_2^{+*} = 0$ for each $\alpha \in \{0.1, 0.33, 0.5, 0.67, 0.9\}$. Therefore, these systems are stochastic $\alpha-$MPSS by Definition (3.1). Also, the above table expresses that for a set of n systems, if $\alpha < \alpha'$,

Table 2
The data set of numerical example.

DMU_j	Input 1	Input 2	Output 1	Output 2
DMU_1	$X_{11} \sim LN(20, 25)$	$X_{21} \sim LN(25, 16)$	$Y_{11} \sim LN(1000, 100)$	$Y_{21} \sim LN(900, 400)$
DMU_2	$X_{12} \sim LN(15, 4)$	$X_{22} \sim LN(23, 18)$	$Y_{12} \sim LN(800, 200)$	$Y_{22} \sim LN(950, 300)$
DMU_3	$X_{13} \sim LN(10, 4)$	$X_{23} \sim LN(9, 9)$	$Y_{13} \sim LN(950, 400)$	$Y_{23} \sim LN(500, 450)$
DMU_4	$X_{14} \sim LN(18, 8)$	$X_{24} \sim LN(10, 8)$	$Y_{14} \sim LN(850, 500)$	$Y_{24} \sim LN(550, 430)$
DMU_5	$X_{15} \sim LN(17, 6)$	$X_{25} \sim LN(18, 7)$	$Y_{15} \sim LN(980, 550)$	$Y_{25} \sim LN(800, 100)$
DMU_6	$X_{16} \sim LN(16, 4)$	$X_{26} \sim LN(19, 15)$	$Y_{16} \sim LN(700, 520)$	$Y_{26} \sim LN(600, 250)$
DMU_7	$X_{17} \sim LN(11, 9)$	$X_{27} \sim LN(20, 14)$	$Y_{17} \sim LN(750, 700)$	$Y_{27} \sim LN(650, 230)$
DMU_8	$X_{18} \sim LN(19, 20)$	$X_{28} \sim LN(17, 4)$	$Y_{18} \sim LN(850, 350)$	$Y_{28} \sim LN(830, 450)$
DMU_9	$X_{19} \sim LN(12, 10)$	$X_{29} \sim LN(15, 17)$	$Y_{19} \sim LN(600, 150)$	$Y_{29} \sim LN(580, 160)$
DMU_{10}	$X_{110} \sim LN(13, 5)$	$X_{210} \sim LN(10, 12)$	$Y_{110} \sim LN(970, 300)$	$Y_{210} \sim LN(560, 400)$
DMU_{11}	$X_{111} \sim LN(16, 6)$	$X_{211} \sim LN(22, 16)$	$Y_{111} \sim LN(780, 110)$	$Y_{211} \sim LN(700, 350)$
DMU_{12}	$X_{112} \sim LN(9, 4)$	$X_{212} \sim LN(8, 3)$	$Y_{112} \sim LN(650, 90)$	$Y_{212} \sim LN(860, 310)$

Table 3
Results of the stochastic $\alpha-$MPSS systems

	α-MPSS	α-MPSS	α-MPSS	α-MPSS	α-MPSS
DMU_j	$\alpha = 0.1$	$\alpha = 0.33$	$\alpha = 0.5$	$\alpha = 0.67$	$\alpha = 0.9$
DMU_1	1	1	1	1	1
DMU_2	1	1	1	1	1
DMU_3	$e^{15.11}$	$e^{26.9}$	e^{34}	$e^{41.1}$	$e^{54.6}$
DMU_4	e^{116}	e^{128}	e^{135}	e^{142}	e^{156}
DMU_5	1	$e^{5.91}$	e^{13}	$e^{20.1}$	$e^{33.6}$
DMU_6	e^{263}	e^{283}	e^{294}	e^{301}	e^{315}
DMU_7	e^{210}	e^{230}	e^{241}	e^{248}	e^{261}
DMU_8	$e^{54.8}$	$e^{70.3}$	$e^{78.6}$	$e^{87.3}$	e^{105}
DMU_9	e^{303}	e^{318}	e^{327}	e^{335}	e^{353}
DMU_{10}	1	$e^{7.9}$	e^{15}	$e^{22.1}$	$e^{35.63}$
DMU_{11}	e^{172}	e^{190}	e^{200}	e^{211}	e^{233}
DMU_{12}	$e^{41.8}$	$e^{63.6}$	e^{75}	$e^{86.4}$	e^{108}

then the number of systems stochastic $\alpha'-$MPSS is less than or equal to the number of systems stochastic $\alpha-$MPSS.

5 Conclusion

The purpose of classic MDEA model is to evaluate the performance of a set of DMUs by considering precise data. These models are very sensitive to measurement errors and data entry errors. Therefore, in real-world scenarios, stochastic models may be better suited for MDEA model when there exists uncertainty associated with the inputs and or outputs of DMUs. To estimate MPSS in the presence of inputs and outputs having log-normal distributions, the input–output orientation model that was introduced by Banker and Maindiratta (1986) in classic MDEA is developed in stochastic MDEA. The deterministic equivalent of stochastic input output oriented MDEA model is obtained and also the concepts of stochastic $\alpha-$MPSS is defined. As an example of system reliability was used to demonstrate the capability of the proposed approach. This example was run in five cases of α and it is observed that the number of systems featured stochastic $\alpha-$MPSS decreases when the value of α increases. In order to further studies, the approach of this research may be extended to some other distributions.

References

[1] Banker, R. D. (1984). Estimating most productive scale size using data envelopment analysis. European Journal of Operational Research, 17, 35–44.

[2] Banker, R. D. (1993). Maximum likelihood, consistency and DEA: statistical foundations. Management Science, 39, 1265–1273.

[3] Banker, R. D., Charnes, A., & Cooper, W. W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science, 30, 1078–1092.

[4] Banker, R. D., Charnes, A., Cooper, W. W., & Schinnar, A. (1981). A bi-extremal principle for frontier estimation and efficiency evaluation. Management Science, 27, 1370–1382.

[5] Banker, R. D., Cooper, W. W., Thrall, R. M., Seiford, L. M., & Zhu, J. (2004). Returns to scale in different DEA models. European Journal of Operational Research, 154, 345–362.

[6] Banker, R. D., & Maindiratta, A. (1986). Piecewise log-linear estimation of efficient production surfaces. Management Science, 32, 126–135.

[7] Cao, X., & Yang, F. (2011). Measuring the performance of internet companies using a two-stage data envelopment analysis model. Enterprise Information Systems, 5, 207–217.

[8] Chang, K. P., & Guh, Y. (1994). Piecewise loglinear frontiers and log efficiency measures. Computers and Operations Research, 22, 1031–1037.

[9] Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2(6), 429–444.

[10] Charnes, A., Cooper, W. W., Seiford, L., & Stutz, J. (1982). A multiplicative model for efficiency analysis. Socio-Economic Planning Sciences, 16, 213–224.

[11] Charnes, A., Cooper, W. W., Seiford, L. M., & Stutz, J. (1983). Invariant multiplicative efficiency and piecewise Cobb-Douglas envelopments. Operations Research Letters, 2, 101–103.

[12] Cook, W. D., & Seiford, L. M. (2009). Equivalence and implementation of alternative methods for determining returns to scale in data envelopment analysis. European Journal of Operational Research, 192, 1–17.

[13] Cooper, W. W., Deng, H., Huang, Z. M., & Li, S. X. (2002). Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. Journal of the Operational Research Society, 53, 1347–1356.

[14] Cooper, W. W., Deng, H., Huang, Z. M., & Li, S. X. (2003). Chance constrained programming approaches to congestion in stochastic data envelopment analysis. European Journal of Operational Research, 155, 231–238.

[15] Cooper, W. W., Huang, Z. M., Lelas, V., Li, S. X., & Olesen, O. B. (1998). Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA. Journal of Productivity Analysis, 9, 53–79.

[16] Cooper, W. W., Seiford, L. M., & Tone, K. (2006). Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software (2nd ed.). New York: Springer.

[17] Cooper, W. W., Thompson, R. G., & Thrall, R. M. (1996). Extensions and new developments in DEA. The Annals of Operations Research, 66, 3–45.

[18] Davoodi, A., Zarepisheh, M., & Zhiani Rezai, H., (2015). The nearest MPSS pattern in data envelopment analysis, Annals of Operations Research, 226, 163–176.

[19] Hollingsworth, B., Dawson, P. J., & Maniadakis, N. (1999). Efficiency measurement of health care: A review of nonparametric methods and applications. Health Care Management Science, 2, 161–172.

[20] Huang, Z. M., & Li, S. X. (1996). Dominance stochastic models in data envelopment analysis. European Journal of Operational Research, 95, 370–403.

[21] Jahanshahloo, G. R., & Khodabakhshi, M. (2003). Using input-output orientation model for determining most productive scale size in DEA. Applied Mathematics and Computation, 146, 849–855.

[22] Khodabakhshi, M. (2009). Estimating most productive scale size with stochastic data in data envelopment analysis. Economic Modelling, 26(5), 968–973.

[23] Land, K. C., Lovell, C. A. K., & Thore, S. (1993). Chance constrained data envelopment analysis. Managerial and Decision Economics, 14, 541–554.

[24] Lee, C.-Y. (2015). Most productive scale size versus demand fulfillment: A solution to the capacity dilemma. European Journal of Operational Research, 1–9.

[25] Mehdiloozad, M., Sahoo, B. K., & Roshdi, I. (2014). Generalized multiplicative directional distance function for efficiency measurement in DEA. European Journal of Operational Research, 3, 679–688.

[26] Menger, K. (1954). The laws of return: a study in meta-economics. In O. Morgenstern (Ed.) Economic activity analysis, part 111, New York: Wiley.

[27] Olesen, O. B., & Petersen, N. C. (1995). Chance constrained efficiency evaluation. Management Science, 41, 442–457.

[28] Seiford, L. M., & Zhu, J. (1998). On piecewise loglinear frontiers and log efficiency measures. Computers and Operations Research, 25, 389–395.

[29] Stapelberg, R. F. (2009). Handbook of reliability, availability, maintainability and safety in engineering design (1nd ed.). London: Springer.

[30] Sueyoshi, T., & Chang, Y. (1989). Efficient algorithms for additive and multiplicative models in data envelopment analysis. Operations Research Letters, 8, 205–213.

[31] Xing, Y., Li, L., Bi, Z., Wilamowska-Korsak, M., & Zhang, L. (2013). Operations research (OR) in service industries: A comprehensive review. Systems Research and Behavioral Science, 30, 300–353.

[32] Zarepisheh, M., Khorram, E., & Jahanshahloo, G. R. (2009). Returns to scale in multiplicative models in data envelopment analysis. Annals of Operations Research, 173, 195–206.

[33] Zarepisheh, M., Khorram, E., & Jahanshahloo, G. R. (2010). Returns to scale in multiplicative models in data envelopment analysis. Annals of Operations Research, 173, 195–206.