Journal of Mathematical Modeling
Vol. -, No. -, 2026, pp. -. Research Article

JMM

Resource allocation and investment analysis with stochastic data

Saeid Ghobadi[†∗], **Khosro Soleimani-Chamkhorami**[‡]

[†]*Department of Mathematics, Kho.C., Islamic Azad University, Khomeinishahr, Iran.*
[†]*Efficiency and Smartization of Energy Systems Research Center, Kho.C., Islamic Azad University, Khomeinishahr, Iran.*
[‡]*Department of Mathematics, Na.C., Islamic Azad University, Najafabad, Iran.*

Email(s): ghobadi.s@iau.ac.ir, kh.soleimani@iau.ac.ir

Abstract. Resource allocation and investment analysis of decision-making units (DMUs) are significant economic and strategic planning management issues. Although various methods have been developed to address these issues based on the inverse data envelopment analysis (DEA), no study has addressed resource allocation and investment analysis in stochastic DEA under efficiency improvement at significance level $\alpha \in [0,1]$. Therefore, this paper studies the resource allocation and investment analysis of DMUs under a deviation from the efficiency frontier due to noise and random errors in the data. The provided approach is formulated using stochastic multiple-objective programming (SMOP) problems. Unlike other proposed approaches, the current paper considers the simultaneous increase and decrease of the various resources/products under stochastic data. Necessary and sufficient conditions are constructed for resources-estimation or products-estimation using stochastic (weak) Pareto solution of SMOP problems at the significance level $\alpha \in [0,1]$. Also, a banking application is utilized to evaluate the performance of the presented theory.

Keywords: DEA, inverse DEA, stochastic data, efficiency, stochastic multiple-objective programming.
AMS Subject Classification 2010: 90C08, 90B50, 90C29.

1 Introduction

Data envelopment analysis (DEA), a useful mathematical programming-based tool for determining the efficiency of decision-making units (DMUs), has been proposed by Charnes et al. [9] for the first time and extended by other researchers. According to the data type, various models have been proposed, including fuzzy, interval, and stochastic models. Since data uncertainty cannot be covered through traditional DEA,

∗Corresponding author
Received: 26 December 2025/ Revised: 13 August 2026/ Accepted: 14 August 2026
DOI: 10.22124/jmm.2026.32637.2965

the uncertainty models were employed. Peykani et al. [42] categorized the uncertain theoretical methods in the uncertain DEA (UDEA) into the following five categories: (i) The fuzzy DEA (FDEA) method; (ii) The bootstrap DEA (BDEA) method; (iii) The imprecise DEA (IDEA) method; (iv) The robust DEA (RDEA) method; and (v) The stochastic DEA (SDEA) method.

Fuzzy DEA approach was employed to simulate vague inputs/output with fuzzy numbers [41]. According to the literature, FDEA is divided into six categories, involving the tolerance, the fuzzy ranking, the α-level based, the possibility, the fuzzy arithmetic, and the fuzzy random/type-2 fuzzy set methods [17]. When the number of DMUs is insufficient, while the estimated efficient frontier differs from the actual one, the BDEA method is applied [45]. The IDEA method was adopted in cases where inputs/outputs contain imprecise information in the form of the interval, ratio bounded, and weak and strong ordinals [53]. As a subset of the UDEA approach, RDEA is a suitable approach to solve optimization problems. This method does not need remarkable historical data and probability distribution function and guarantees the solution feasibility for all possible amounts of uncertain parameters in the chosen convex uncertainty set [42].

The SDEA is another UDEA technique applied for evaluating the efficiency of DMUs under uncertain data. Since inaccurate data is a fundamental challenge for decision-makers in the real world, they may be described with stochastic variables. A variety of data characteristics can be derived when dealing with stochastic parameters, assuming prognostication events. Predicting the upcoming efficiencies is an essential advantage for working with stochastic data in DEA. Bolós et al. [7] proposed chance-constrained directional DEA models that generalize stochastic radial DEA and introduced a stochastic version of the generalized Farrell measure. Arabi et al. [4] proposed a stochastic network DEA model integrated with chance-constrained programming to evaluate the circular economy performance of European countries, explicitly accounting for data uncertainty and the interdependence between economic production and waste management subsystems. Bolós et al. [8] introduced the SDEAR package, providing a comprehensive computational framework for implementing chance-constrained SDEA models. Deng and Li [13] proposed a fuzzy stochastic DEA model that incorporates the input–output structure of decision-making units to achieve a more comprehensive efficiency ranking. In the DEA literature, various theoretical and practical aspects of the SDEA models have been revealed by several scholars, some of which are presented in Table 1.

Wei et al. [50] introduced the concept of InvDEA. They utilized multiple-objective programming (MOP) for estimating output/input levels under increasing inputs/outputs and preserve the efficiency index. Afterward, various researchers developed the InvDEA models. Jahanshahloo et al. [33, 34] converted multiple-objective linear programming (MOLP) into single-objective linear programming based on the decision-maker's preferences. Recently, Emrouznejad et al. [16] presented a comprehensive review of InvDEA, discussing its origins, methodological developments, applications, and future research directions. Also, Emrouznejad and Amin [15] highlighted the growing importance of InvDEA in performance assessment and emphasized its recent methodological advances and expanding applications. More recently, Ghobadi et al. [27] introduced generalized inverse DEA-R frameworks for restructuring evaluation, further expanding the methodological developments and practical applications of InvDEA. Several new InvDEA works are presented in Table 2. In this article, the InvDEA approach is utilized to improve the efficiency index for stochastic data.

Resource management means developing the organization's resources effectively. These resources involve financial resources, inventory, human skills, production resources, information technology (IT), and natural resources. Resource allocation aims to plan, manage, and assign resources properly to attain

Table 1: Selected studies in SDEA

Reference	Short description
[38]	This work develops DEA models under random disturbances.
[32]	In this study, the IDEA and inverse IDEA models are used to address revenue-setting problems.
[43]	This paper applies SDEA in contexts where input prices and capital adjustment costs are subject to variation.
[35]	The research suggested a method for ranking stochastic DMUs based on the reliability of their efficiency.
[5]	The study proposed a linear programming model based on symmetric error structure for inputs and outputs.
[44]	The work suggested a data generation algorithm for a multivariate, cross-sectional, nonparametric stochastic frontier.
[48]	The focus of this research is the development of DEA-based chance-constrained programming.
[11]	The reference constructed a stochastic production possibility set (PPS) to propose a SDEA method.
[12]	The authors here propose a DEA model for random inputs and outputs and introduce a stochastic efficient DMU.
[37]	This study estimates the most productive scale size in SDEA.
[6]	The work proposes a ranking system with random data.
[31]	This research employs the coefficient of variation to rank DMUs with stochastic data.
[30]	They extended the centralized resource allocation method in the presence of random data.
[39]	This study compared chance-constrained DEA, StoNED, and bootstrap methods for efficiency evaluation under uncertainty.
[1]	The work integrated SDEA with GARMA for dynamic efficiency analysis and forecasting of undesirable outputs.

Table 2: Landmark studies in InvDEA

Reference	Short description
[3]	This research estimates potential gains from bank mergers using a two-stage InvDEA model.
[52]	In the work, the authors introduced the InvDEA model based on non-radial DEA.
[47]	InvDEA models are designed to maintain cost and revenue efficiency.
[46]	A ranking system based on InvDEA is proposed in this study.
[18]	This article studies the allocation of CO2 emission quotas under various considerations.
[40]	This study proposes a dynamic inverse network DEA model for supply chain sustainability.
[2]	The issue of merging units based on InvDEA in the presence of negative data has been examined.
[25]	An InvDEA framework for merger target setting with fuzzy data is developed in this study.
[20]	This study estimates cost efficiency under available data prices using InvDEA.
[21]	This paper proposes an InvDEA model for cost and revenue efficiency without price information.
[22]	This study generalizes InvDEA through the directional distance function.
[19, 51]	In these studies, InvDEA models have been employed to address the merger of DMUs.

the organization's strategic purposes. It improves the performance of the manager's work. In strategic planning, resource allocation employs available resources to attain future purposes. In economics, public finance includes the following three broad fields: macroeconomic stabilization, the distribution of income and wealth, and resource allocation. Various resource allocation strategies have been focused on deriving the conditions under which certain resource allocation mechanisms yield Pareto efficient outcomes. Scholars established various approaches to analyze efficiency and resource allocation. The mentioned approaches are divided into parametric and nonparametric approaches. The two commonly used approaches employed for efficiency analysis, known as DEA, are the parametric stochastic frontier generation function and the nonparametric mathematical programming approaches.

Furthermore, investment analysis is a broad term for various approaches of assessing investments, industry sectors, and economic trends. The investment analysis determines the performance of an investment and suitability for a specific investor. The main parameters in investment analysis involve the suitable entry price, the expected time horizon to hold an investment, and the total investment role in the portfolio. Investment analysis includes assessing a comprehensive investment policy based on the performed thought process, the person's requirements and the current financial condition, the portfolio performance, and the modification or adjustment time. Investment analysis approaches include bottom-up, top-down, fundamental, and technical. Bottom-up investing is an investment method that analyzes individual stocks and de-emphasizes the importance of macroeconomic and market cycles. Bottom-up investing necessitates investors first to verify microeconomic variables. As an investment analysis method, top-down investing concentrates on the macro parameters of the economy like gross domestic product (GDP), employment, taxation, and interest rates. Top-down investing prioritizes macroeconomic, national, or market-level variables. Fundamental analysis (FA) is a conventional approach that performs the analysis through determining the investment's Fair market value. Accordingly, the investor decides whether to purchase the company stock. It is yet another great and efficient investment analysis approach. The technical analysis approach is utilized to indicate and distinguish the business chances by considering the stock market statistics.

According to the DEA literature, InvDEA idea has been employed in the resource allocation and investment analysis of DMUs from theoretical and practical perspectives. To the best of the authors' knowledge, none of these studies account for deviations from the frontier due to noise and random error in the data, with the exception of Ghomi et al. [28]. They extended the InvDEA problems to a stochastic analytical framework. In other words, they answered the following important question: If the efficiency score remains unchanged at the significance level $\alpha \in [0,1]$, but the inputs/outputs are increased, to what extent should the outputs/inputs of the unit be increased? Ghomi et al. [28] employed stochastic multiple-objective programming (SMOP) tools to estimate the inputs/outputs. Their proposed models cannot be employed under the arbitrary variations in input/output levels and the efficiency improvement a the significance level α. The inputs/outputs change indicates that some of the inputs/outputs would be increased while some others decreased or remain the same. Therefore, the current work presents a novel InvDEA technique for the resource allocation and investment analysis of units under stochastic data and the arbitrary variations in resources/products and the efficiency improvement at the significance level α. This is the main contribution of the current paper. The proposed approach is formulated based on the SMOP model. A decision-making environment is considered where a decision-maker controls the resources and products of a set of units. The InvDEA and SMOP are combined to allow a decision-maker to incorporate preference information under stochastic data. Necessary and sufficient conditions are derived for input resources estimation or output products estimation utilizing stochastic Pareto and weak Pareto solutions at significance level $\alpha \in [0,1]$ of SMOP problems. The efficiency of the proposed theory is verified in a banking sector application.

Organization of the current article is given as the following. Section 2 discusses some preliminaries on SDEA. The paper's main results are presented in Sections 3 and 4. In Section 3, a stochastic InvDEA model is constructed for allocating resources based on the SMOP framework. Also, Section 4 performs an investment analysis under improvement efficiency in the presence of stochastic data. A numerical example is presented in Section 5 to explain the applicability of this proposed method. Section 6 presents the conclusions, limitations, and directions for future research.

2 Stochastic DEA

In this section, the fundamental principles of SDEA relevant to this article are outlined. For this purpose, suppose a set of n DMUs, $\{DMU_j : j \in J = \{1,2,\ldots,n\}\}$, which DMU_j consume positive input resources vector $\tilde{X}_j = (\tilde{x}_{1j}, \tilde{x}_{2j}, \ldots, \tilde{x}_{mj})^t$ to produce positive output products vector $\tilde{Y}_j = (\tilde{y}_{1j}, \tilde{y}_{2j}, \ldots, \tilde{y}_{sj})^t$. Suppose that the input resources and output products of DMU_j, $j \in J$, are normal-stochastic numbers as follows:

$$\tilde{x}_{ij} \sim N(x_{ij}, \sigma_{ij}^2), \quad \forall i \in I = \{1, 2, \ldots, m\},$$
$$\tilde{y}_{rj} \sim N(y_{rj}, \psi_{rj}^2), \quad \forall r \in O = \{1, 2, \ldots, s\},$$

where x_{ij} and σ_{ij}^2 for all $i \in I$ and $j \in J$, are mean and variance of input resources, respectively, y_{rj} and ψ_{rj}^2 for all $r \in O$ and $j \in J$, are mean and variance of output products, respectively. The unit under assessment is called DMU_o, $o \in J$. The following input-oriented SDEA model is applied for assessing the relative efficiency of DMU_o at the significance level α [10] under constant returns to scale (CRS) assumption in the production technology:

$$\begin{aligned} \theta_o(\alpha) = \min \quad & \theta \\ s.t. \quad & P\{\sum_{j=1}^{n} \mu_j \tilde{X}_j \leq \ \theta \tilde{X}_o\} \geq 1 - \alpha, \\ & P\{\sum_{j=1}^{n} \mu_j \tilde{Y}_j \geq \tilde{Y}_o\} \geq 1 - \alpha, \\ & \mu = (\mu_1, \ldots, \mu_n) \geq 0, \end{aligned} \tag{1}$$

where, P means “Probability” and $\alpha \in [0,\ 1]$ is a predetermined is a level of error and $(\mu, \theta) \in \mathbb{R}^n \times \mathbb{R}$ is variables vector. Also, the efficiency score of DMU_o at significance level α is denoted $eff(\tilde{X}_o, \tilde{Y}_o)(\alpha) = \theta_o(\alpha)$. Let Φ be the cumulative standard normal distribution function, then the stochastic model (1) change into the following non-linear model [10]:

$$\begin{aligned} \theta_o(\alpha) = \min \quad & \theta \\ s.t. \quad & \sum_{j=1}^{n} \mu_j x_{ij} - \Phi^{-1}(\alpha) \sigma_i^I(\mu, \theta) \leq \ \theta x_{io}, \quad \forall i \in I, \\ & \sum_{j=1}^{n} \mu_j y_{rj} + \Phi^{-1}(\alpha) \sigma_r^O(\mu) \geq y_{ro}, \quad \forall r \in O, \\ & \mu = (\mu_1, \ldots, \mu_n) \geq 0, \end{aligned} \tag{2}$$

where

$$\left(\sigma_i^I(\mu, \theta)\right)^2 = \sum_{j=1}^{n} \sum_{k=1}^{n} \mu_j \mu_k Cov(\tilde{x}_{ij}, \tilde{x}_{ik}) + \theta^2 Var(\tilde{x}_{io}) - 2\theta \sum_{j=1}^{n} \mu_j Cov(\tilde{x}_{ij}, \tilde{x}_{io}), \quad \forall i \in I,$$

$$\left(\sigma_r^O(\mu)\right)^2 = \sum_{j=1}^{n} \sum_{k=1}^{n} \mu_j \mu_k Cov(\tilde{y}_{rj}, \tilde{y}_{rk}) + Var(\tilde{y}_{ro}) - 2 \sum_{j=1}^{n} \mu_j Cov(\tilde{y}_{rj}, \tilde{y}_{ro}), \quad \forall r \in O.$$

If all inputs and outputs are uncorrelated, then

$$\left(\sigma_i^I(\mu, \theta)\right)^2 = \sum_{j=1}^{n} \mu_j^2 Var(\tilde{x}_{ij})^2 + \theta Var(\tilde{x}_{io})^2 (\theta - 2\mu_o), \quad \forall i \in I,$$

$$(\sigma_r^O(\mu))^2 = \sum_{j=1}^{n} \mu_j^2 Var(\tilde{y}_{rj}) + Var(\tilde{y}_{ro})^2(1-2\mu_o), \quad \forall r \in O.$$

The uncorrelated case is presented here for expository purposes; however, model (2) includes covariance terms $(\sigma_i^I(\mu,\theta))^2$ and $(\sigma_r^O(\mu))^2$ that capture arbitrary correlations among inputs and among outputs.

Similarly, the output-oriented of stochastic DEA model at significance level α under CRS assumption in the production technology, is as follows [10]:

$$\begin{aligned} \varphi_o(\alpha) = \max \quad & \varphi \\ s.t. \quad & P\{\sum_{j=1}^{n} \mu_j \tilde{X}_j \leq \tilde{X}_o\} \geq 1-\alpha, \\ & P\{\sum_{j=1}^{n} \mu_j \tilde{Y}_j \geq \varphi \tilde{Y}_o\} \geq 1-\alpha, \\ & \mu = (\mu_1, \ldots, \mu_n) \geq 0. \end{aligned} \tag{3}$$

In this model, $(\mu,\varphi) \in \mathbb{R}^n \times \mathbb{R}$ is the vector of decision variables. Also, the efficiency score of DMU_o at significance level α is denoted $eff(\tilde{X}_o, \tilde{Y}_o)(\alpha) = \varphi_o(\alpha)$. Moreover, if Φ is cumulative standard distribution function, the stochastic model (3) change into model (4):

$$\begin{aligned} \varphi_o(\alpha) = \max \quad & \varphi \\ s.t. \quad & \sum_{j=1}^{n} \mu_j x_{ij} - \Phi^{-1}(\alpha)\sigma_i^I(\mu) \leq x_{io}, \quad \forall i \in I, \\ & \sum_{j=1}^{n} \mu_j y_{rj} + \Phi^{-1}(\alpha)\sigma_r^O(\mu,\varphi) \geq \varphi y_{ro}, \quad \forall r \in O, \\ & \mu = (\mu_1, \ldots, \mu_n) \geq 0, \end{aligned} \tag{4}$$

where

$$(\sigma_i^I(\mu))^2 = \sum_{j=1}^{n}\sum_{k=1}^{n} \mu_j \mu_k Cov(\tilde{x}_{ij}, \tilde{x}_{ik}) + Var(\tilde{x}_{io}) - 2\sum_{j=1}^{n} \mu_j Cov(\tilde{x}_{ij}, \tilde{x}_{io}), \quad \forall i \in I,$$

$$(\sigma_r^O(\mu,\varphi))^2 = \sum_{j=1}^{n}\sum_{k=1}^{n} \mu_j \mu_k Cov(\tilde{y}_{rj}, \tilde{y}_{rk}) + \varphi^2 Var(\tilde{y}_{ro}) - 2\varphi\sum_{j=1}^{n} \mu_j Cov(\tilde{y}_{rj}, \tilde{y}_{ro}), \quad \forall r \in O.$$

Definition 1. *The unit under evaluation, DMU_o, is called stochastic input (output)-oriented weakly efficient at significance level α if $\theta_o(\alpha) = 1$ ($\varphi_o(\alpha) = 1$).*

Theorem 1. *([10]) i) Models (2) and (4) are feasible for all significance level α.*
ii) If $\alpha \leq 0.5$, then $0 < \theta_o(\alpha) \leq 1$ and $\varphi_o(\alpha) \geq 1$.
iii) If $\alpha \leq \alpha'$, then $\theta_o(\alpha') \leq \theta_o(\alpha)$ and $\varphi_o(\alpha') \geq \varphi_o(\alpha)$.

Remark 1. *It is worth noting that, for DMU_o, the input- and output-oriented efficiency scores obtained from Models (2) and (4) at significance level α are reciprocals of each other; that is, $\theta_o(\alpha) = \frac{1}{\varphi_o(\alpha)}$.*

We conclude this section by presenting the conditions and the mathematical derivation required to transform models (2) and (4) into their equivalent LP problems [5]. Assume that the input resources and output products of the DMU_j, $j = 1,\dots,n$ have asymmetric error structure as following structure [5]:

$$\begin{aligned} \tilde{x}_{ij} &= x_{ij} + a_{ij}\tilde{\varepsilon}_{ij}, \quad \forall i \in I, \\ \tilde{y}_{rj} &= y_{rj} + b_{rj}\tilde{\xi}_{rj}, \quad \forall r \in O, \end{aligned} \tag{5}$$

where a_{ij} and b_{rj} are nonnegative real values, and $\tilde{\varepsilon}_{ij}$ and $\tilde{\xi}_{rj}$ are random variables with normal distributions, $\tilde{\varepsilon}_{ij} \sim N(0,\bar{\sigma}^2)$ and $\tilde{\xi}_{rj} \sim N(0,\bar{\sigma}^2)$. Therefore, $\tilde{\varepsilon}ij$ and $\tilde{\xi}rj$ denote the input and output errors with respect to their corresponding mean values, respectively. Because the normal distribution is symmetric around its mean, the error structure described by (5) is termed the symmetric error structure. The following relationships are then obtained from (5):

$$\begin{aligned} \tilde{x}_{ij} &\sim N(x_{ij},\bar{\sigma}^2 a_{ij}^2), \quad \forall i \in I, \\ \tilde{y}_{rj} &\sim N(y_{rj},\bar{\sigma}^2 b_{rj}^2), \quad \forall r \in O. \end{aligned} \tag{6}$$

Any random variable following a normal distribution can be represented by a symmetric error structure. Furthermore, assume that the ith inputs are mutually uncorrelated across all DMUs. Similarly, let the rth outputs be mutually uncorrelated. Hence, for all $j \neq k$,

$$\begin{aligned} Cov(\tilde{\varepsilon}_{ij},\tilde{\varepsilon}_{ik}) &= 0, \quad \forall i \in I, \\ Cov(\tilde{\xi}_{rj},\tilde{\xi}_{rk}) &= 0, \quad \forall r \in O. \end{aligned} \tag{7}$$

Hence, it follows from (5) and (7) that the same error term may be assumed for all DMUs. Specifically, $\tilde{\varepsilon}_{ij} = \tilde{\varepsilon}_i$ and $\tilde{\xi}_{rj} = \tilde{\xi}_r$ for all $i \in I$ and $r \in O$. Consider now the ith input constraint of model (1):

$$P\{\sum_{j=1}^{n} \mu_j \tilde{x}_{ij} \leq \ \theta \tilde{x}_{io}\} \geq 1-\alpha. \tag{8}$$

Let $\tilde{h}_i = \sum_{j=1}^{n} \mu_j \tilde{x}_{ij} - \theta \tilde{x}_{io}$. It follows from (5) and (7) that:

$$\tilde{h}_i = (\sum_{j=1}^{n} \mu_j x_{ij} - \theta x_{io}) + \tilde{\varepsilon}_i(\sum_{j=1}^{n} \mu_j a_{ij} - \theta a_{io}). \tag{9}$$

Therefore,

$$\tilde{h}_i \sim N((\sum_{j=1}^{n} \mu_j x_{ij} - \theta x_{io}),\bar{\sigma}^2(\sum_{j=1}^{n} \mu_j a_{ij} - \theta a_{io})^2). \tag{10}$$

Using the above expression together with the properties of the normal distribution, the stochastic constraint (8) is equivalent to the following deterministic constraint:

$$\sum_{j=1}^{n} \mu_j x_{ij} - \Phi^{-1}(\alpha)\bar{\sigma}\left|\sum_{j=1}^{n} \mu_j a_{ij} - \theta a_{io}\right| \leq \ \theta x_{io}, \quad \forall i \in I. \tag{11}$$

Similarly, the rth output constraint of model (1) can be converted into the following deterministic equivalent:

$$\sum_{j=1}^{n}\mu_j y_{rj}+\Phi^{-1}(\alpha)\bar{\sigma}\left|\sum_{j=1}^{n}\mu_j b_{rj}-b_{ro}\right|\geq y_{ro},\quad \forall r\in O. \tag{12}$$

Consequently, using (11) and (12), model (1) can be reformulated as the following deterministic equivalent:

$$\begin{aligned}\theta_o(\alpha)=\min\quad & \theta\\ s.t.\quad & \sum_{j=1}^{n}\mu_j x_{ij}-\Phi^{-1}(\alpha)\bar{\sigma}\left|\sum_{j=1}^{n}\mu_j a_{ij}-\theta a_{io}\right|\leq\ \theta x_{io},\quad \forall i\in I,\\ & \sum_{j=1}^{n}\mu_j y_{rj}+\Phi^{-1}(\alpha)\bar{\sigma}\left|\sum_{j=1}^{n}\mu_j b_{rj}-b_{ro}\right|\geq y_{ro},\quad \forall r\in O,\\ & \mu=(\mu_1,\ldots,\mu_n)\geq 0.\end{aligned} \tag{13}$$

Although model (13) is a nonlinear programming model, it can be reformulated as an LP model using the following transformations:

$$\begin{aligned}&\left|\sum_{j=1}^{n}\mu_j a_{ij}-\theta a_{io}\right|=p_i^{+}+p_i^{-},\quad \forall i\in I,\\ &\sum_{j=1}^{n}\mu_j a_{ij}-\theta a_{io}=p_i^{+}-p_i^{-},\quad \forall i\in I,\\ &p_i^{+}p_i^{-}=0,p_i^{+}\geq 0,p_i^{-}\geq 0,\quad \forall i\in I,\end{aligned} \tag{14}$$

and

$$\begin{aligned}&\left|\sum_{j=1}^{n}\mu_j b_{rj}-b_{ro}\right|=q_r^{+}+q_r^{-},\quad \forall r\in O,\\ &\sum_{j=1}^{n}\mu_j b_{rj}-b_{ro}=q_r^{+}-q_r^{-},\quad \forall r\in O,\\ &q_r^{+}q_r^{-}=0,q_r^{+}\geq 0,q_r^{-}\geq 0,\quad \forall r\in O.\end{aligned} \tag{15}$$

Substituting the above transformation into model (13) yields:

$$\begin{aligned}\theta_o(\alpha)=\min\quad & \theta\\ s.t.\quad & \sum_{j=1}^{n}\mu_j x_{ij}-\Phi^{-1}(\alpha)\bar{\sigma}(p_i^{+}+p_i^{-})\leq\ \theta x_{io},\quad \forall i\in I,\\ & \sum_{j=1}^{n}\mu_j y_{rj}+\Phi^{-1}(\alpha)\bar{\sigma}(q_r^{+}+q_r^{-})\geq y_{ro},\quad \forall r\in O,\\ & \sum_{j=1}^{n}\mu_j a_{ij}-\theta a_{io}=p_i^{+}-p_i^{-},\quad \forall i\in I,\\ & \sum_{j=1}^{n}\mu_j b_{rj}-b_{ro}=q_r^{+}-q_r^{-},\quad \forall r\in O,\end{aligned} \tag{16}$$

$$p_i^+ p_i^- = 0, q_r^+ q_r^- = 0, \forall i \in I, \forall r \in O,$$
$$\mu_j \geq 0, p_i^+ \geq 0, p_i^- \geq 0, q_r^+ \geq 0, q_r^- \geq 0, \forall j \in J, \forall i \in I, \forall r \in O.$$

Model (16) is nonlinear due to the complementarity constraints $p_i^+ p_i^- = 0$ and $q_r^+ q_r^- = 0$. It is well known that every LP problem with an optimal solution has at least one basic optimal solution. Omitting the complementarity constraints $p_i^+ p_i^- = 0$ and $q_r^+ q_r^- = 0$ transforms model (16) into an LP model. Furthermore, every optimal basic solution of the resulting model satisfies the property that, for each $i \in I$, at least one of p_i^+ and p_i^- is zero, and for each $r \in O$, at least one of q_r^+ and q_r^- is zero. Therefore, every optimal basic solution of the resulting linear model satisfies the complementarity constraints $p_i^+ p_i^- = 0$ and $q_r^+ q_r^- = 0$. Consequently, when the model is solved using an algorithm that yields an optimal basic solution (e.g.,, the Simplex method), the complementarity constraints $p_i^+ p_i^- = 0$ and $q_r^+ q_r^- = 0$ may be safely omitted, since they are satisfied automatically by every optimal basic solution. Therefore, the linear deterministic equivalent of model (1) is given by:

$$\begin{aligned} \theta_o(\alpha) = \min \quad & \theta \\ s.t. \quad & \sum_{j=1}^{n} \mu_j x_{ij} - \Phi^{-1}(\alpha)\bar{\sigma}(p_i^+ + p_i^-) \leq \theta x_{io}, \quad \forall i \in I, \\ & \sum_{j=1}^{n} \mu_j y_{rj} + \Phi^{-1}(\alpha)\bar{\sigma}(q_r^+ + q_r^-) \geq y_{ro}, \quad \forall r \in O, \\ & \sum_{j=1}^{n} \mu_j a_{ij} - \theta a_{io} = p_i^+ - p_i^-, \quad \forall i \in I, \\ & \sum_{j=1}^{n} \mu_j b_{rj} - b_{ro} = q_r^+ - q_r^-, \quad \forall r \in O, \\ & \mu_j \geq 0, p_i^+ \geq 0, p_i^- \geq 0, q_r^+ \geq 0, q_r^- \geq 0, \forall j \in J, \forall i \in I, \forall r \in O. \end{aligned} \tag{17}$$

The decision variables of model (17) are $\mu_j, p_i^+, p_i^-, q_r^+, q_r^-$, and θ for all relevant indices. Similarly, model (3) can be reformulated as the following LP model:

$$\begin{aligned} \varphi_o(\alpha) = \max \quad & \varphi \\ s.t. \quad & \sum_{j=1}^{n} \mu_j x_{ij} - \Phi^{-1}(\alpha)\bar{\sigma}(p_i^+ + p_i^-) \leq x_{io}, \quad \forall i \in I, \\ & \sum_{j=1}^{n} \mu_j y_{rj} + \Phi^{-1}(\alpha)\bar{\sigma}(q_r^+ + q_r^-) \geq \varphi y_{ro}, \quad \forall r \in O, \\ & \sum_{j=1}^{n} \mu_j a_{ij} - a_{io} = p_i^+ - p_i^-, \quad \forall i \in I, \\ & \sum_{j=1}^{n} \mu_j b_{rj} - \varphi b_{ro} = q_r^+ - q_r^-, \quad \forall r \in O, \\ & \mu_j \geq 0, p_i^+ \geq 0, p_i^- \geq 0, q_r^+ \geq 0, q_r^- \geq 0, \forall j \in J, \forall i \in I, \forall r \in O. \end{aligned} \tag{18}$$

The decision variables of model (18) are $\mu_j, p_i^+, p_i^-, q_r^+, q_r^-$, and φ for all relevant indices.

Remark 2. *The proposed stochastic DEA models are consistent with their deterministic counterparts. In particular, when $\alpha = 0.5$, we have $\Phi^{-1}(0.5) = 0$, and therefore the stochastic terms in Models (17)* *and (18) vanish. Likewise, as $\bar{\sigma} \longrightarrow 0$, the stochastic perturbation terms disappear. Hence, the proposed stochastic InvDEA models developed in Sections 3 and 4 reduce to their corresponding deterministic InvDEA models.*

3 Resource allocation under improvement of efficiency

Effective resource allocation ensures that each unit's inputs are directed to the highest-value use, ensuring profitability in competitive markets. This is particularly important when units deviate from the efficiency frontier due to noise and random errors in the data. Therefore, this section is devoted to extending the following question, provided by [29], to a stochastic setting.

Question: Among a set of units, to what extent should the input resources of the unit change if the efficiency score of the unit remains unchanged or improves it to the amount η-percent, yet the output products change?

Remark 3. *When $\eta = 0$, the proposed models represent the efficiency-preservation case. Compared with the stochastic InvDEA models of Ghomi et al. [28], the proposed formulation additionally permits arbitrary changes in the input/output levels, including increases, decreases, and unchanged values. Hence, the proposed models constitute a generalization of the approach in [28].*

Addressing the above question enables decision-makers to make more informed decisions aimed at developing DMUs. Consequently, by selecting an appropriate strategy for expanding a unit, decision-makers can undertake the necessary operational actions. From a managerial standpoint, resolving this question is crucial for preserving strategic control during transitional periods. Market fluctuations, shifts in demand patterns, and changes in product mixes are inevitable; nevertheless, management must ensure that such dynamics do not undermine operational efficiency. In fact, a precise understanding of how inputs should be adjusted in response to output changes-whether to maintain the current efficiency level or to achieve a targeted η-percent improvement-allows decision-makers to establish realistic production targets, negotiate budgets with greater confidence, and identify underperforming units requiring corrective intervention. This understanding transforms efficiency analysis from a passive reporting mechanism into an active framework for resource reallocation, thereby ensuring that every operational change is anchored in data-driven accountability.

The above question was addressed using MOP formulations by [29, 50]. Different theoretical and practical perspectives of the resource allocation have been studied in the DEA literature, such as setting revenue target [14], investment analysis [49], and resource allocation [18] under inter-temporal dependence of data [23, 24, 36], restructuring DMUs [26]. Nevertheless, the constructed models in the mentioned works could not solve the issue of resource allocation in the presence of stochastic data. Therefore, this section attempts to estimate resource allocation under efficiency improvement in the presence of stochastic data.

Suppose the output products of $DMU_o(\tilde{X}_o, \tilde{Y}_o)$ are changed from $\tilde{Y}_o$ to $\tilde{\beta}_o = \tilde{Y}_o + \Delta\tilde{Y}_o$, where $\Delta\tilde{Y}_o \neq 0$. We need to estimate the input resources vector $\tilde{\alpha}_o = (\tilde{\alpha}_{1o}, \tilde{\alpha}_{2o}, \cdots, \tilde{\alpha}_{mo})$ provided that the efficiency index of DMU_o at the significance level α, maintains its current efficiency level, that is $\theta_o(\alpha)$, or improves it to the amount η-percent. In fact, $\tilde{\alpha}_o = \tilde{X}_o + \Delta\tilde{X}_o$, in which $\Delta\tilde{X}_o \neq 0$.

Suppose $DMU_{new}(\tilde{\alpha}_o, \tilde{\beta}_o)$ represents the unit obtained after changing the input resources and output products of $DMU_o(\tilde{X}_o, \tilde{Y}_o)$. The following model measures the efficiency score of $DMU_{new}(\tilde{\alpha}_o, \tilde{\beta}_o)$ at the significance level α:

$$
\begin{aligned}
\theta_{new}(\alpha) = \min \quad & \theta \\
s.t. \quad & P\{\sum_{j=1}^{n} \mu_j \tilde{X}_j + \mu_{new} \tilde{\alpha}_o \le \theta \tilde{\alpha}_o\} \ge 1-\alpha, \\
& P\{\sum_{j=1}^{n} \mu_j \tilde{Y}_j + \mu_{new} \tilde{\beta}_o \ge \tilde{\beta}_o\} \ge 1-\alpha, \\
& \mu^+ = (\mu_1, \ldots, \mu_n, \mu_{new}) \ge 0.
\end{aligned}
\tag{19}
$$

In the above model, $(\mu^+, \theta) \in \mathbb{R}^{n+1} \times \mathbb{R}$ is variables vector. It is worth noting that $\mu^+ = (\mu_1, \ldots, \mu_n, \mu_{new})$ denotes the extended intensity vector obtained by augmenting the original intensity vector $\mu = (\mu_1, \ldots, \mu_n)$ with the additional intensity variable μ_{new} corresponding to $DMU_{new}(\tilde{\alpha}_o, \tilde{\beta}_o)$. Thus, μ is used for the original technology, whereas μ^+ is used for the expanded technology. More precisely, in model (19), the intensity variable μ_{new} corresponds to the newly formed unit $DMU_{new}(\tilde{\alpha}_o, \tilde{\beta}_o)$. At first glance, allowing $\mu_{new} > 0$ might appear circular, because the new unit's inputs $\tilde{\alpha}_o$ and outputs $\tilde{\beta}_o$ are themselves decision variables to be estimated. However, this is not circular reasoning for the following reasons: (i)The variable μ_{new} represents the weight assigned to the new unit in the production possibility set (PPS) after its creation. In InvDEA, the new unit is not assumed to be part of the original technology; rather, the expanded technology is defined as the convex hull of the original n DMUs plus the new unit. The intensity μ_{new} quantifies how much the new unit contributes to the convex combinations that define the efficiency frontier of the expanded technology. (ii) Theorems 2 and 3 establish that any stochastic weak Pareto solution of SMOP (21) guarantees that the efficiency target is achieved without requiring $\mu_{new} > 0$. Indeed, in the proof of Theorem 2, we construct a feasible solution to model (21) with $\mu_{new} = 0$ (see Eq. (25) and the subsequent argument). This demonstrates that the desired efficiency improvement can be achieved even when the new unit is not used as a reference for itself. The case $\mu_{new} > 0$ only arises if the new unit's optimal intensity in the expanded frontier is positive, which is a natural outcome of the convexity of the PPS.

Remark 4. *If all input sources and output products are assumed to have a symmetric error structure, with independent normally distributed random errors sharing a common standard deviation $\bar{\sigma}$, then the stochastic model (19) can be transformed into the linear programming model (20) using the same transformation procedure as that employed for converting Model (1) into Model (17):*

$$
\begin{aligned}
\theta_{new}(\alpha) = \min \quad & \theta \\
s.t. \quad & \sum_{j=1}^{n} \mu_j x_{ij} + \mu_{new} \alpha_{io} - \Phi^{-1}(\alpha) \bar{\sigma} (p_i^+ + p_i^-) \le \theta \alpha_{io}, \quad \forall i \in I, \\
& \sum_{j=1}^{n} \mu_j y_{rj} + \mu_{new} \beta_{ro} + \Phi^{-1}(\alpha) \bar{\sigma} (q_r^+ + q_r^-) \ge \beta_{ro}, \quad \forall r \in O, \\
& \sum_{j=1}^{n} \mu_j a_{ij} + \mu_{new} a_{io} - \theta a_{io} = p_i^+ - p_i^-, \quad \forall i \in I, \\
& \sum_{j=1}^{n} \mu_j b_{rj} + \mu_{new} b_{ro} - b_{ro} = q_r^+ - q_r^-, \quad \forall r \in O, \\
& \mu_{new} \ge 0, \mu_j \ge 0, p_i^+ \ge 0, p_i^- \ge 0, q_r^+ \ge 0, q_r^- \ge 0, \forall j \in J, \forall i \in I, \forall r \in O.
\end{aligned}
\tag{20}
$$

The decision variables of model (20) are $\mu_{new}, \mu_j, p_i^+, p_i^-, q_r^+, q_r^-$, and θ for all relevant indices.

The following definition is established based on the optimal solutions of models (1) and (19) at the significance level α.

Definition 2. *Suppose that $\theta_o(\alpha)$ and $\theta_{new}(\alpha)$ are the optimal values of problems (1) and (19), respectively. Then*

i) If $\theta_o(\alpha) = \theta_{new}(\alpha)$, then it is said that the efficiency score of DMU_o remains unchanged a the significance level α, i.e.,, $eff(\tilde{\alpha}_o, \tilde{\beta}_o)(\alpha) = eff(\tilde{X}_o, \tilde{Y}_o)(\alpha)$[1].

ii) If $\theta_{new}(\alpha) = (1 + \frac{\eta}{100})\theta_o(\alpha)$, then it is said that the amount of improvement of the efficiency of DMU_o is η-percent of the $\theta_o(\alpha)$ at the significance level α, i.e.,, $eff(\tilde{\alpha}_o, \tilde{\beta}_o) = (1 + \frac{\eta}{100})eff(\tilde{X}_o, \tilde{Y}_o)$.

To answer the question, i.e., to estimate the input resources, along the lines of [29] SMOP is used. We introduce the SMOP in the next subsections.

3.1 General idea of the SMOP formulation

The classic InvDEA problem asks: if a unit changes its outputs (or inputs), how should its inputs (or outputs) be adjusted such that its efficiency score does not deteriorate? In the stochastic setting, where data are subject to random noise, this question becomes more challenging because efficiency is defined through probabilistic constraints.

We consider DMU_o which a manager plans to modify its output vector from $\tilde{\mathbf{Y}}_o$ to a new target $\tilde{\beta}_o$. The outputs may increase for some products and decrease for others, reflecting strategic reallocations. The manager seeks to achieve a certain improvement in the unit's efficiency score - for instance, raising it from its current level $\theta_o(\alpha)$ to $(1+\eta/100)\theta_o(\alpha)$- by adjusting the input vector to a new value $\tilde{\alpha}_o$. The key question is: what is the smallest possible new input vector (in the Pareto sense) that makes this efficiency improvement attainable with prescribed probability $1-\alpha$?

To answer this question, we formulate an SMOP model. The model simultaneously minimizes each component of the new input vector, because lower resource consumption is always preferable from a managerial perspective. The constraints capture two essential requirements with a confidence level $1-\alpha$:

- **Efficiency condition:** The existing set of units, when combined with appropriate intensity weights, should be able to produce the new input vector at a scaled level that reflects the desired efficiency improvement.

- **Output feasibility:** The same combination of existing units must be able to produce at least the target outputs $\tilde{\beta}_o$.

These conditions are imposed as chance constraints, because the data are stochastic. The SMOP framework allows us to seek solutions that are not dominated in all input components - i.e.,, stochastic Pareto or weak Pareto solutions. A solution is considered 'good' if there is no other feasible solution that reduces some input component(s) without increasing others, while still satisfying the probabilistic constraints.

In essence, the SMOP formulation translates the resource allocation problem under uncertainty into a multi-objective optimization problem where the trade-offs among different inputs are explicitly considered. The resulting Pareto frontier provides the manager with a set of efficient input vectors, each

[1] Hereafter, we use the notation " eff " instead of " efficiency " for simplicity.

corresponding to a different balance between reducing one resource versus another, while guaranteeing the desired efficiency improvement with high confidence.

The following subsections formalize this idea mathematically, define the stochastic Pareto optimality concepts, and provide necessary and sufficient conditions for the estimated input vector to achieve the exact efficiency target.

3.2 Proposed SMOP model

We consider a set of n decision-making units DMU$_j$ $(j = 1,\dots,n)$ with stochastic input vectors $\tilde{\mathbf{X}}_j = (\tilde{x}_{1j},\dots,\tilde{x}_{mj})^t$ and stochastic output vectors $\tilde{\mathbf{Y}}_j = (\tilde{y}_{1j},\dots,\tilde{y}_{sj})^t$. All random variables are assumed to be normally distributed, and the inputs and outputs are allowed to be correlated. The production technology is assumed to exhibit CRS. For a specific unit under consideration, DMU$_o$, let $\theta_o(\alpha)$ denote its input-oriented efficiency score at significance level $\alpha \in [0,1]$, obtained from model (1) (or its deterministic equivalent (2)). Suppose the output vector of DMU$_o$ is changed from $\tilde{\mathbf{Y}}_o$ to a given target vector $\tilde{\boldsymbol{\beta}}_o = (\tilde{\beta}_{1o},\dots,\tilde{\beta}_{so})$ with $\Delta\tilde{\mathbf{Y}}_o = \tilde{\boldsymbol{\beta}}_o - \tilde{\mathbf{Y}}_o \neq \mathbf{0}$. The decision-maker wants to estimate a new input vector $\tilde{\boldsymbol{\alpha}}_o = (\tilde{\alpha}_{1o},\dots,\tilde{\alpha}_{mo})$ such that after the change, the efficiency score of DMU$_o$ becomes $(1+\eta/100)\,\theta_o(\alpha)$, where η is a given improvement percentage.

To obtain $\tilde{\boldsymbol{\alpha}}_o$, we formulate the following SMOP model.

$$
\begin{aligned}
\min \quad & (\tilde{\alpha}_{1o},\tilde{\alpha}_{2o},\dots,\tilde{\alpha}_{mo}) \\
s.t. \quad & P\{\sum_{j=1}^{n}\mu_j\tilde{X}_j \leq (1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o\} \geq 1-\alpha, \\
& P\{\sum_{j=1}^{n}\mu_j\tilde{Y}_j \geq \tilde{\beta}_o\} \geq 1-\alpha, \\
& \mu = (\mu_1,\dots,\mu_n) \geq 0.
\end{aligned}
\tag{21}
$$

The rationale is to simultaneously minimize the expected values of the new inputs while ensuring that the original set of units can produce at least the target outputs and that the efficiency condition is satisfied with a given probability. The value $\theta_o(\alpha)$ is the current efficiency score of DMU$_o$ (a fixed constant), and η is a user-specified improvement percentage satisfying $0 \leq \eta \leq \frac{1-\theta_o(\alpha)}{\theta_o(\alpha)} \times 100$. If $\eta = 0$, the efficiency remains unchanged; if η reaches its upper bound, the unit becomes fully efficient (i.e.,, $\theta_{\text{new}}(\alpha) = 1$).

The first chance constraint guarantees that, with probability at least $1-\alpha$, the convex combination of the existing input vectors (using intensities μ) does not exceed the scaled new input vector of DMU$_o$. The second chance constraint ensures that the same convex combination can produce at least the target outputs $\tilde{\beta}_o$ with the same probability level.

Since the objective is to minimize all components of $\tilde{\boldsymbol{\alpha}}_o$ simultaneously, (21) is an SMOP. Its solutions are defined in the sense of stochastic (weak) Pareto optimality (see Definitions 3 and 4). The structure of (21) allows the decision-maker to explore trade-offs among the new input levels; any reduction in one input component may require an increase in another to maintain the probabilistic efficiency condition.

Under the assumption that all random inputs and outputs are normally distributed (or have a symmetric error structure), the chance constraints in (21) can be transformed into deterministic nonlinear constraints using the cumulative distribution function Φ and covariances. After this transformation, the SMOP problem becomes a deterministic multiple-objective programming problem, which can be solved by standard methods (e.g.,, weighting or ε-constraint techniques) to generate a set of Pareto-efficient input vectors $\tilde{\alpha}_o$. Theorems 2 and 3 (presented in the following subsection) establish the equivalence between the stochastic weak Pareto solutions of (21) and the desired efficiency target.

Remark 5. *If the efficiency score of DMU_o is $\theta_o(\alpha)$ at the significance level α, then η must be $0 \leq \eta \leq \frac{1-\theta_o(\alpha)}{\theta_o(\alpha)} * 100$. If $\eta = 0$, then DMU_o, with respect to other units, maintains its current efficiency score at the significance level α. If $\eta = \frac{1-\theta_o(\alpha)}{\theta_o(\alpha)} * 100$, then DMU_o will be efficient at the significance level α.*

Remark 6. *Under the assumptions stated in Remark 4, the SMOP model (21) can be transformed into Model (22) by following the same procedure used to transform Model (1) into Model (17).*

$$
\begin{aligned}
\min \quad & (\alpha_{1o}, \alpha_{2o}, \ldots, \alpha_{mo}) \\
s.t. \quad & \sum_{j=1}^{n} \mu_j x_{ij} - \Phi^{-1}(\alpha)\bar{\sigma}(p_i^+ + p_i^-) \leq (1 + \frac{\eta}{100})\theta_o(\alpha)\alpha_{io}, \quad \forall i \in I, \\
& \sum_{j=1}^{n} \mu_j y_{rj} + \Phi^{-1}(\alpha)\bar{\sigma}(q_r^+ + q_r^-) \geq \beta_{ro}, \quad \forall r \in O, \\
& \sum_{j=1}^{n} \mu_j a_{ij} - (1 + \frac{\eta}{100})\theta_o(\alpha) a_{io} = p_i^+ - p_i^-, \quad \forall i \in I, \\
& \sum_{j=1}^{n} \mu_j b_{rj} - b_{ro} = q_r^+ - q_r^-, \quad \forall r \in O, \\
& \mu_j \geq 0, p_i^+ \geq 0, p_i^- \geq 0, q_r^+ \geq 0, q_r^- \geq 0, \forall j \in J, \forall i \in I, \forall r \in O.
\end{aligned}
\tag{22}
$$

The decision variables of model (22) are $\mu_j, p_i^+, p_i^-, q_r^+, q_r^-, a_{io}$, and α_{io} for all relevant indices. It is worth noting that $\Phi^{-1}(\alpha)$, the inverse cumulative distribution function of the standard normal distribution, satisfies $\Phi^{-1}(\alpha) < 0$ for $\alpha < 0.5$. Consequently, in the first deterministic constraint, the term $-\Phi^{-1}(\alpha)\bar{\sigma}(p_i^+ + p_i^-)$ becomes positive, whereas in the second deterministic constraint, the term $+\Phi^{-1}(\alpha)\bar{\sigma}(q_r^+ + q_r^-)$ becomes negative. In both cases, the deterministic reformulation becomes more restrictive, thereby preserving the prescribed confidence level of the original chance constraints. From a managerial perspective, this conservative reformulation reflects a risk-aware decision policy under uncertainty. By incorporating chance constraints, the model introduces appropriate safety margins for both inputs and outputs, thereby reducing the risk of infeasible or unreliable efficiency assessments arising from random variations in the data. Consequently, the resulting efficiency scores are more robust and provide decision-makers with a more reliable basis for benchmarking and performance evaluation in uncertain environments.

Due to the special structure of SMOP (21), we define a new optimality (efficiency) concept for this problem at the significance level α.

Definition 3. *A feasible solution $\Delta = (\mu^*, \tilde{\alpha}_{1o}^*, \tilde{\alpha}_{2o}^*, \ldots, \tilde{\alpha}_{mo}^*)$ is called a stochastic Pareto solution in the level of significance α to SMOP (21) if there does not exist another feasible solution $\Gamma = (\mu, \tilde{\alpha}_{1o}, \tilde{\alpha}_{2o}, \ldots, \tilde{\alpha}_{mo})$ such that*

$$P\{\tilde{\alpha}_{io} - \tilde{\alpha}_{io}^* \leq 0\} \geq 1 - \alpha \text{ for each} \quad i = 1, 2, \ldots, m,$$

$$P\{\tilde{\alpha}_{io} - \tilde{\alpha}_{io}^* \leq -\xi\} \geq 1 - \alpha \text{ for some} \quad i = 1, 2, \ldots, m,$$

where ξ a non-Archimedian infinitesimal [2].

Definition 4. *A feasible solution $\Delta = (\mu^*, \tilde{\alpha}^*_{1o}, \tilde{\alpha}^*_{2o}, ..., \tilde{\alpha}^*_{mo})$ is called a stochastic weak Pareto solution at the significance level α to SMOP (21) if there does not exist another feasible solution $\Gamma = (\mu, \tilde{\alpha}_{1o}, \tilde{\alpha}_{2o}, ..., \tilde{\alpha}_{mo})$ such that*

$$P\{\tilde{\alpha}_{io} - \tilde{\alpha}^*_{io} \leq -\xi\} \geq 1-\alpha \quad \textit{for each} \quad i = 1,2,\ldots,m,$$

where ξ a non-Archimedian infinitesimal.

Theorem 2 shows how the above SMOP can be used to estimate the input resources.

Theorem 2. *Suppose that $\Delta = (\mu^*, \tilde{\alpha}^*_o)$ is a stochastic weak Pareto solution at the significance level α to model (21). Then*

$$eff(\tilde{\alpha}^*_o, \tilde{\beta}_o)(\alpha) = (1+\frac{\eta}{100})eff(\tilde{X}_o, \tilde{Y}_o)(\alpha).$$

Proof. To prove the theorem, we should show that $\theta_{new}(\alpha) = (1+\frac{\eta}{100})\theta_o(\alpha)$, considering $\tilde{\alpha}^*_o = (\tilde{\alpha}^*_{1o}, \tilde{\alpha}^*_{2o}, ..., \tilde{\alpha}^*_{mo})$ in model (19). Feasibility of Δ for SMOP (21), implies

$$P\{\sum_{j=1}^{n} \mu^*_j \tilde{X}_j \leq (1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}^*_o\} \geq 1-\alpha, \tag{23}$$

$$P\{\sum_{j=1}^{n} \mu^*_j \tilde{Y}_j \geq \tilde{\beta}_o\} \geq 1-\alpha, \tag{24}$$

$$\mu^* = (\mu^*_1, ..., \mu^*_n) \geq 0. \tag{25}$$

Let $\bar{\mu} = (\bar{\mu}_1, \ldots, \bar{\mu}_n, \bar{\mu}_{new})$, in which $\bar{\mu}_j = \mu^*_j$ for each $j = 1, \ldots, n$ and $\bar{\mu}_{new} = 0$. According to (23), (24), and (25), $(\bar{\mu},\ (1+\frac{\eta}{100})\theta_o(\alpha))$ is a feasible solution to problem (19) (considering $\tilde{\alpha}^*_o$ in stead of $\tilde{\alpha}_o$). Therefore, $\theta_{new}(\alpha) \leq (1+\frac{\eta}{100})\theta_o(\alpha)$.

Let $(\mu^{+*} = (\mu^{+*}_1, \ldots, \mu^{+*}_n, \mu^{+*}_{new}), \theta_{new}(\alpha) = \theta^{+*})$ be an optimal solution to problem (19), we have

$$P\{\sum_{j=1}^{n} \mu^{+*}_j \tilde{X}_j + \mu^{+*}_{new}\tilde{\alpha}^*_o \leq \theta^{+*}\tilde{\alpha}^*_o\} \geq 1-\alpha, \tag{26}$$

$$P\{\sum_{j=1}^{n} \mu^{+*}_j \tilde{Y}_j + \mu^{+*}_{new}\tilde{\beta}_o \geq \tilde{\beta}_o\} \geq 1-\alpha, \tag{27}$$

$$\mu^{+*} = (\mu^{+*}_1, \ldots, \mu^{+*}_n, \mu^{+*}_{new}) \geq 0. \tag{28}$$

By Remark 5, we have $0 < (1+\frac{\eta}{100})\theta_o(\alpha) \leq 1$. Therefore, with regard to inequality (23), we obtain

$$P\{\sum_{j=1}^{n} \mu^*_j \tilde{X}_j \leq \tilde{\alpha}^*_o\} \geq 1-\alpha. \tag{29}$$

[2]Here, ξ is a non-Archimedean infinitesimal used to distinguish strict Pareto improvement from weak Pareto improvement, consistent with standard multi-objective programming conventions.

Using (26) and (29), the following result is obtained:

$$P\{\sum_{j=1}^{n}\mu_j^{+*}\tilde{X}_j+\mu_{new}^{+*}(\sum_{j=1}^{n}\mu_j^{*}\tilde{X}_j)\leq\theta^{+*}\tilde{\alpha}_o^*\}\geq 1-\alpha, \tag{30}$$

then

$$P\{\sum_{j=1}^{n}(\mu_j^{+*}+\mu_j^{*}\mu_{new}^{+*})\tilde{X}_j\leq\theta^{+*}\tilde{\alpha}_o^*\}\geq 1-\alpha. \tag{31}$$

In a similar manner, by Eqs. (24) and (27) we get

$$P\{\sum_{j=1}^{n}(\mu_j^{+*}+\mu_j^{*}\mu_{new}^{+*})\tilde{Y}_j\geq\tilde{\beta}_o\}\geq 1-\alpha. \tag{32}$$

Set $\overline{\mu}_j:=\mu_j^{+*}+\mu_j^{*}\mu_{new}^{+*}$ for each $j=1,2,\ldots,n$. It is easily seen that

$$\overline{\mu}=(\overline{\mu}_j,\ldots,\overline{\mu}_n)\geq 0. \tag{33}$$

By definition of $\overline{\mu}$, we can write inequalities (31) and (32) as follows:

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j\leq\theta^{+*}\tilde{\alpha}_o^*\}\geq 1-\alpha, \tag{34}$$

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{Y}_j\geq\tilde{\beta}_o\}\geq 1-\alpha. \tag{35}$$

Now, by contradiction assume that $\theta_{new}(\alpha)=\theta^{+*}<(1+\frac{\eta}{100})\theta_o(\alpha)$. By Eq. (34) and contradiction assume, we obtain

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j<(1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o^*\}\geq 1-\alpha, \tag{36}$$

then

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o^*<0\}\geq 1-\alpha. \tag{37}$$

It is obvious that there exists non-negative vector $\varepsilon_o=(\varepsilon_{1o},\varepsilon_{2o},...,\varepsilon_{mo})$ such that

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o^*<0\}=1-\alpha+\varepsilon_o, \tag{38}$$

since the left-hand side is strictly negative with probability at least $1-\alpha$, there exists a positive vector $S_o=(s_{1o},s_{2o},...,s_{mo})$ such that the inequality holds with strict negative margin. Then

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o^*\leq -S_o\}=1-\alpha. \tag{39}$$

Clearly, there exists a positive vector $\overline{S}_o=(\overline{s}_{1o},\overline{s}_{2o},\ldots,\overline{s}_{mo})$ such that $S_o=(1+\frac{\eta}{100})\theta_o(\alpha)\overline{S}_o$. Then, by Eq. (39), we get

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o^*\leq -(1+\frac{\eta}{100})\theta_o(\alpha)\overline{S}_o\}=1-\alpha. \tag{40}$$

Therefore,

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)(\tilde{\alpha}_o^*-\overline{S}_o)\leq 0\}=1-\alpha. \tag{41}$$

Let $\overline{\tilde{\alpha}}_o := \tilde{\alpha}_o^*-\overline{S}_o$. By (41), we have

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)\overline{\tilde{\alpha}}_o\leq 0\}\geq 1-\alpha. \tag{42}$$

Also $-\overline{k} =: \max\{-\overline{s}_{io},\ \ i=1,2,...,m\}$, with regard to Eqs. (33), (35), and (42), it is obvious that $(\overline{\mu},\overline{\tilde{\alpha}}_o)$ is a feasible solution to problem (21), where $P\{\overline{\tilde{\alpha}}_o-\tilde{\alpha}_o^*\leq -\overline{k}\}\geq 1-\alpha$. But this is impossible because Δ is a stochastic weak Pareto solution at the significance level α to SMOP (21).□

Remark 7. *If $P\{|\tilde{\alpha}_o^*-X_o|<\xi\}$ where ξ is a non-Archimedian infinitesimal, and there exists at least $l\in I$ such that $P\{\tilde{\beta}_{lo}-\tilde{y}_{lo}>\varepsilon\}\geq 1-\alpha$, then $\tilde{\beta}_{lo}-\tilde{y}_{lo}$ indicates the lack-output product amount in lth output product component of the DMU$_o$. In other words, the decision-maker can preserve the efficiency score of the DMU$_o$ while the output products increase from $\tilde{Y}_o$ to $\tilde{\beta}_o$ without the input resources increase from $\tilde{X}_o$.*

Theorem 3 give converse version of Theorem 2.

Theorem 3. *Let $\Lambda=(\overline{\mu},\overline{\tilde{\alpha}_o})$ be a feasible solution to SMOP (21) such that $eff(\overline{\tilde{\alpha}}_o,\tilde{\beta}_o)(\alpha)=(1+\frac{\eta}{100})eff(\tilde{X}_o,\tilde{Y}_o)(\alpha)$ where $0\leq\eta\leq\frac{1-\theta_o^*}{\theta_o^*}*100$. Then Λ is a stochastic weak Pareto solution at the significance level α to SMOP (21).*

Proof. If Λ is not be a stochastic weak Pareto solution at the significance level α to problem (21), then there would exist another feasible solution of (21), $\Delta=(\mu,\tilde{\alpha}_o=(\tilde{\alpha}_{1o},\tilde{\alpha}_{2o},...,\tilde{\alpha}_{mo}))$ in which

$$P\{\tilde{\alpha}_o-\overline{\tilde{\alpha}}_o\leq -\xi\}\geq 1-\alpha, \tag{43}$$

where ξ is a non-Archimedian infinitesimal. Feasibility of Δ for SMOP (21), implies

$$P\{\sum_{j=1}^{n}\mu_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o\leq 0\}\geq 1-\alpha, \tag{44}$$

$$P\{\sum_{j=1}^{n}\mu_j\tilde{Y}_j\geq\tilde{\beta}_o\}\geq 1-\alpha, \tag{45}$$

$$\mu=(\mu_1,...,\mu_n)\geq 0. \tag{46}$$

By Eqs. (43) and (44), we get

$$P\{\sum_{j=1}^{n}\mu_j\tilde{X}_j-(1+\frac{\eta}{100})\theta_o(\alpha)(\overline{\tilde{\alpha}}_o-\varepsilon)\leq 0\}\geq 1-\alpha. \tag{47}$$

Therefore, there exists positive scalar $0<k_o<1$ such that

$$P\{\sum_{j=1}^{n}\mu_j\tilde{X}_j-(k_o(1+\frac{\eta}{100})\theta_o(\alpha))\overline{\tilde{\alpha}}_o\leq 0\}\geq 1-\alpha. \tag{48}$$

According to Eqs. (45), (46), and (48), the vector $((\mu,\mu_{new}=0), k_o(1+\frac{\eta}{100})\theta_o(\alpha))$ is a feasible solution to model (19). The value of the objective function of model (19) at this feasible point is equal to $k_o(1+\frac{\eta}{100})\theta_o(\alpha)$. Therefore,

$$eff(\bar{\tilde{\alpha}}_o,\tilde{\beta}_o)(\alpha)=\theta_{new}(\alpha)\leq k_o(1+\frac{\eta}{100})\theta_o(\alpha)<(1+\frac{\eta}{100})\theta_o(\alpha)=eff(\tilde{X}_o,\tilde{Y}_o)(\alpha).$$

This contradicts the assumption and completes the proof.□

4 Investment analysis under improvement of efficiency

This section is devoted to extending the following question, provided by [50], to a stochastic setting.

Question: Among a set of units, to what extent should the output products of the unit change if the efficiency score of the unit remains unchanged or improves it to the amount η-percent, yet the input resources change?

Suppose the input resources of $DMU_o(\tilde{X}_o,\tilde{Y}_o)$ are changed from $\tilde{X}_o$ to $\tilde{\alpha}_o=\tilde{X}_o+\Delta\tilde{X}_o$, where $\Delta\tilde{X}_o\neq 0$. We need to estimate the output products vector $\tilde{\beta}_o=(\tilde{\beta}_{1o},\tilde{\beta}_{2o},\cdots,\tilde{\beta}_{so})$ provided that the efficiency index of DMU$_o$ at the significance level α, maintains its current efficiency level, that is $\varphi_o(\alpha)$, or improves it to the amount η-percent. In fact, $\tilde{\beta}_o=\tilde{Y}_o+\Delta\tilde{Y}_o$, in which $\Delta\tilde{Y}_o\neq 0$.

Suppose $DMU_{new}(\tilde{\alpha}_o,\tilde{\beta}_o)$ represents the unit obtained after changing the inputs and outputs of $DMU_o(\tilde{X}_o,\tilde{Y}_o)$. The following output-oriented model measures the efficiency score of $DMU_{new}(\tilde{\alpha}_o,\tilde{\beta}_o)$ in the level of significance α:

$$\begin{aligned}
\varphi_{new}(\alpha)=\max\quad & \varphi\\
s.t.\quad & P\{\sum_{j=1}^{n}\mu_j\tilde{X}_j+\mu_{new}\tilde{\alpha}_o\leq\tilde{\alpha}_o\}\geq 1-\alpha,\\
& P\{\sum_{j=1}^{n}\mu_j\tilde{Y}_j+\mu_{new}\tilde{\beta}_o\geq\varphi\tilde{\beta}_o\}\geq 1-\alpha,\\
& \mu^{+}=(\mu_1,\ldots,\mu_n,\mu_{new})\geq 0.
\end{aligned}\tag{49}$$

In the above model, $(\mu^{+},\varphi)\in\mathbb{R}^{n+1}\times\mathbb{R}$ is variables vector.

Remark 8. *Under the assumptions stated in Remark 4, the stochastic model (49) can be transformed into Model (50) by following the same procedure used to transform Model (4) into Model (18).*

$$\begin{aligned}
\varphi_{new}(\alpha)=\max\quad & \varphi\\
s.t.\quad & \sum_{j=1}^{n}\mu_j x_{ij}+\mu_{new}\alpha_{io}-\Phi^{-1}(\alpha)\bar{\sigma}(p_i^{+}+p_i^{-})\leq\ \alpha_{io},\quad\forall i\in I,\\
& \sum_{j=1}^{n}\mu_j y_{rj}+\mu_{new}\beta_{ro}+\Phi^{-1}(\alpha)\bar{\sigma}(q_r^{+}+q_r^{-})\geq\varphi\beta_{ro},\quad\forall r\in O,\\
& \sum_{j=1}^{n}\mu_j a_{ij}+\mu_{new}a_{io}-a_{io}=p_i^{+}-p_i^{-},\quad\forall i\in I,\\
& \sum_{j=1}^{n}\mu_j b_{rj}+\mu_{new}b_{ro}-\varphi b_{ro}=q_r^{+}-q_r^{-},\quad\forall r\in O,\\
& \mu_{new}\geq 0,\mu_j\geq 0,p_i^{+}\geq 0,p_i^{-}\geq 0,q_r^{+}\geq 0,q_r^{-}\geq 0,\forall i\in I,\forall r\in O.
\end{aligned}\tag{50}$$

The decision variables of model (50) are $\mu_{new}, \mu_j, p_i^+, p_i^-, q_r^+, q_r^-$, and φ for all relevant indices.

The following definition is established based on the optimal solutions of models (3) and (49) at the significance level α.

Definition 5. *Suppose that $\varphi_o(\alpha)$ and $\varphi_{new}(\alpha)$ are the optimal values of problems (3) and (49), respectively. Then*

i) If $\varphi_o(\alpha) = \varphi_{new}(\alpha)$, then it is said that the efficiency score of DMU_o remains unchanged at the significance level α, i.e.,, $eff(\tilde{\alpha}_o, \tilde{\beta}_o)(\alpha) = eff(\tilde{X}_o, \tilde{Y}_o)(\alpha)$.

ii) If $\varphi_{new}(\alpha) = (1 - \frac{\eta}{100})\varphi_o(\alpha)$, then it is said that the amount of improvement of the efficiency of DMU_o is η-percent of the $\varphi_o(\alpha)$ at the significance level α, i.e.,, $eff(\tilde{\alpha}_o, \tilde{\beta}_o) = (1 - \frac{\eta}{100})eff(\tilde{X}_o, \tilde{Y}_o)$.

To answer the question, i.e., to estimate the output products, along the lines of [50], we consider the following SMOP model:

$$
\begin{aligned}
\max \quad & (\tilde{\beta}_{1o}, \tilde{\beta}_{2o}, \ldots, \tilde{\beta}_{so}) \\
s.t. \quad & P\{\sum_{j=1}^{n} \mu_j \tilde{X}_j \leq \tilde{\alpha}_o\} \geq 1 - \alpha, \\
& P\{\sum_{j=1}^{n} \mu_j \tilde{Y}_j \geq (1 - \frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o\} \geq 1 - \alpha, \\
& \mu = (\mu_1, \ldots, \mu_n) \geq 0,
\end{aligned}
\tag{51}
$$

where$\varphi_o(\alpha)$ value is fixed and denote the optimal value of φ variable in model (3). Also, resources of $\tilde{\alpha}_o$ are given and fixed.

Remark 9. *If the efficiency score of DMU_o is $\varphi_o(\alpha)$ at the significance level α, then η must be $0 \leq \eta \leq \frac{\varphi_o(\alpha)-1}{\varphi_o(\alpha)} * 100$. If $\eta = 0$, then DMU_o, with respect to other units, maintains its current efficiency score. If $\eta = \frac{\varphi_o(\alpha)-1}{\varphi_o(\alpha)} * 100$, then DMU_o will be efficient at the significance level α.*

Remark 10. *Under the assumptions stated in Remark 4, the SMOP model (51) can be transformed into Model (52) by following the same procedure used to transform Model (4) into Model (17):*

$$
\begin{aligned}
\max \quad & (\beta_{1o}, \beta_{2o}, \ldots, \beta_{so}) \\
s.t. \quad & \sum_{j=1}^{n} \mu_j x_{ij} - \Phi^{-1}(\alpha)\bar{\sigma}(p_i^+ + p_i^-) \leq \alpha_{io}, \quad \forall i \in I, \\
& \sum_{j=1}^{n} \mu_j y_{rj} + \Phi^{-1}(\alpha)\bar{\sigma}(q_r^+ + q_r^-) \geq (1 - \frac{\eta}{100})\varphi_o(\alpha)\beta_{ro}, \quad \forall r \in O, \\
& \sum_{j=1}^{n} \mu_j a_{ij} - a_{io} = p_i^+ - p_i^-, \quad \forall i \in I, \\
& \sum_{j=1}^{n} \mu_j b_{rj} - (1 - \frac{\eta}{100})\varphi_o b_{ro} = q_r^+ - q_r^-, \quad \forall r \in O, \\
& \mu_j \geq 0, p_i^+ \geq 0, p_i^- \geq 0, q_r^+ \geq 0, q_r^- \geq 0, \forall j \in J, \forall i \in I, \forall r \in O.
\end{aligned}
\tag{52}
$$

The decision variables of model (52) are $\mu_j, p_i^+, p_i^-, q_r^+, q_r^-, b_{ro}$ and β_{ro} for all relevant indices. It should be noted that $\Phi^{-1}(\alpha)$, representing the inverse cumulative distribution function of the standard normal distribution, is negative whenever $\alpha < 0.5$. Accordingly, in the first deterministic constraint, the expression $-\Phi^{-1}(\alpha)\bar{\sigma}(p_i^+ + p_i^-)$ contributes a positive adjustment, whereas in the second deterministic constraint, $+\Phi^{-1}(\alpha)\bar{\sigma}(q_r^+ + q_r^-)$ contributes a negative adjustment. Despite these opposite signs, both effects tighten the deterministic constraints, resulting in a more conservative reformulation that maintains the prescribed confidence level of the original chance-constrained model. In both cases, the deterministic reformulation becomes more restrictive, thereby preserving the prescribed confidence level of the original chance constraints. From a managerial perspective, this conservative reformulation reflects a risk-aware decision policy under uncertainty. By introducing appropriate safety margins for both inputs and outputs, the model reduces the likelihood of unreliable output-target estimations caused by random variations in the data. Consequently, the estimated output targets are more robust and provide decision-makers with a more reliable basis for planning performance improvement under uncertain operating conditions.

Theorem 4 shows how the above SMOP can be used for output products estimation.

Theorem 4. *Suppose that $\Delta = (\mu^*, \tilde{\beta}_o^*)$ is a stochastic weak Pareto solution at the significance level α to Model (51). Then,*

$$eff(\tilde{\alpha}_o, \tilde{\beta}_o^*)(\alpha) = (1 - \frac{\eta}{100}) eff(\tilde{X}_o, \tilde{Y}_o)(\alpha).$$

Proof. To prove the theorem, we should show that $\varphi_o(\alpha) = (1 - \frac{\eta}{100})\varphi_{new}(\alpha)$, considering $\tilde{\beta}_o^* = (\tilde{\beta}_{1o}^*, \tilde{\beta}_{2o}^*, ..., \tilde{\beta}_{so}^*)$ in model (49). Feasibility of Δ for SMOP (51), implies

$$P\{\sum_{j=1}^{n} \mu_j^* \tilde{X}_j \leq \tilde{\alpha}_o\} \geq 1 - \alpha, \tag{53}$$

$$P\{\sum_{j=1}^{n} \mu_j^* \tilde{Y}_j \geq (1 - \frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o^*\} \geq 1 - \alpha, \tag{54}$$

$$\mu^* = (\mu_1^*, ..., \mu_n^*) \geq 0. \tag{55}$$

According to Eqs. (53), (54), and (55), the vector $(\mu = (\mu^*, 0), \varphi = (1 - \frac{\eta}{100})\varphi_o(\alpha))$ is a feasible solution to model (49) (setting $\tilde{\beta}_o^*$ in stead of $\tilde{\beta}_o$). Therefore, $\varphi_{new}(\alpha) \geq (1 - \frac{\eta}{100})\varphi_o(\alpha)$.

Suppose that $(\mu^{+*}, \mu_{new}^{+*}, \varphi_{new}(\alpha) = \varphi^{+*})$ is an optimal solution of model (49), we have

$$P\{\sum_{j=1}^{n} \mu_j^{+*} \tilde{X}_j + \mu_{new}^{+*} \tilde{\alpha}_o \leq \tilde{\alpha}_o\} \geq 1 - \alpha, \tag{56}$$

$$P\{\sum_{j=1}^{n} \mu_j^{+*} \tilde{Y}_j + \mu_{new}^{+*} \tilde{\beta}_o^* \geq \varphi^{+*} \tilde{\beta}_o^*\} \geq 1 - \alpha, \tag{57}$$

$$\mu^{+*} = (\mu_1^{+*}, ..., \mu_n^{+*}, \mu_{new}^{+*}) \geq 0. \tag{58}$$

Considering Eqs. (53) and (56), we have

$$P\{\sum_{j=1}^{n} \mu_j^{+*} \tilde{X}_j + \mu_{new}^{+*}(\sum_{j=1}^{n} \mu_j^* \tilde{X}_j) \leq \tilde{\alpha}_o\} \geq 1 - \alpha, \tag{59}$$

then,

$$P\{\sum_{j=1}^{n}(\mu_j^{+*}+\mu_{new}^{+*}\mu_j^*)\tilde{X}_j)\leq\tilde{\alpha}_o\}\geq 1-\alpha. \tag{60}$$

Let $\overline{\mu}_j=\mu_j^{+*}+\mu_{new}^{+*}\mu_j^*$ for each $j=1,2,\ldots,n$, and write (60) as

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{X}_j\leq\tilde{\alpha}_o\}\geq 1-\alpha. \tag{61}$$

Also, it is obvious that

$$\overline{\mu}=(\overline{\mu}_1,...,\overline{\mu}_n)\geq 0. \tag{62}$$

Since $(1-\frac{\eta}{100})\varphi_o(\alpha)\geq 1$ and Eq. (54), we get

$$P\{\sum_{j=1}^{n}\mu_j^*\tilde{Y}_j\geq\tilde{\beta}_o^*\}\geq 1-\alpha. \tag{63}$$

By contradiction assume that $\varphi_{new}(\alpha)=\varphi^{+*}>(1-\frac{\eta}{100})\varphi_o(\alpha)$. By (57), we get

$$P\{\sum_{j=1}^{n}\mu_j^{+*}\tilde{Y}_j+\mu_{new}^{+*}\tilde{\beta}_o^*>(1-\frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o^*\}\geq 1-\alpha. \tag{64}$$

By (63) and (64), we obtain

$$P\{\sum_{j=1}^{n}\mu_j^{+*}\tilde{Y}_j+\mu_{new}^{+*}(\sum_{j=1}^{n}\mu_j^*\tilde{Y}_j)>(1-\frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o^*\}\geq 1-\alpha, \tag{65}$$

then

$$P\{\sum_{j=1}^{n}(\mu_j^{+*}+\mu_{new}^{+*}(\sum_{j=1}^{n}\mu_j^*))\tilde{Y}_j>(1-\frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o^*\}\geq 1-\alpha. \tag{66}$$

According to definition of $\overline{\mu}_j$ for j=$j=1,2,...,n$, and Eq. (66), we have

$$P\{\sum_{j=1}^{n}\overline{\mu}_j\tilde{Y}_j>(1-\frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o^*\}\geq 1-\alpha, \tag{67}$$

therefore,

$$P\{(1-\frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o^*-\sum_{j=1}^{n}\overline{\mu}_j\tilde{Y}_j)<0\}\geq 1-\alpha. \tag{68}$$

It is obvious that there exists non-negative vector $\varepsilon_o=(\varepsilon_{1o},\varepsilon_{2o},...,\varepsilon_{so})$ such that

$$P\{(1-\frac{\eta}{100})\varphi_o\tilde{\beta}_o^*-\sum_{j=1}^{n}\overline{\mu}_j\tilde{Y}_j)<0\}=1-\alpha+\varepsilon_o, \tag{69}$$

therefore, there exists positive vector $S_o=(s_{1o},s_{2o},...,s_{so})$ such that

$$P\{(1-\frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o^*-\sum_{j=1}^{n}\overline{\mu}_j\tilde{Y}_j)\leq -S_o\}=1-\alpha. \tag{70}$$

Clearly, there exists positive vector $\bar{S}_o = (\bar{s}_{1o}, \bar{s}_{2o}, ..., \bar{s}_{so})$ such that $S_o = (1 - \frac{\eta}{100})\varphi_o(\alpha)\bar{S}_o$. Defining

$$-\bar{k} = \max\{-\bar{s}_{ro},\ \ r = 1, 2, ..., s\}, \tag{71}$$

we have $\bar{k} > 0$. Now, for each $r \in O$, define $\overline{\tilde{\beta}}_{ro} := \tilde{\beta}^*_{ro} + \bar{k}$. By Eq. (70), we get

$$P\{\sum_{j=1}^{n} \overline{\mu}_j \tilde{Y}_j \geq (1 - \frac{\eta}{100})\varphi_o \overline{\tilde{\beta}}_o\} \geq 1 - \alpha. \tag{72}$$

According to Eqs. (61), (62), and (72), $(\overline{\mu}, \overline{\tilde{\beta}_o})$ is a feasible solution to problem (51) where

$$P\{\overline{\tilde{\beta}_{ro}} - \tilde{\beta}^*_{ro} \geq \bar{k}\} \geq 1 - \alpha \qquad \text{for each } r = 1, 2, \ldots, s.$$

But this is impossible because Δ is a stochastic weak Pareto solution at the significance level α to SMOP (51).□

Theorem 5 give converse version of Theorem 4.

Theorem 5. *Let $\Lambda = (\overline{\mu}, \overline{\tilde{\beta}_o} = (\overline{\tilde{\beta}_{1o}}, \overline{\tilde{\beta}_{2o}}, ..., \overline{\tilde{\beta}_{so}}))$ be a feasible solution to SMOP (51) such that $eff(\alpha_o, \overline{\tilde{\beta}_o}) = (1 - \frac{\eta}{100})eff(\tilde{X}_o, \tilde{Y}_o)$ where $0 \leq \eta \leq \frac{\varphi_o(\alpha)-1}{\varphi_o(\alpha)} * 100$. Then Λ is a stochastic weak Pareto solution at the significance level α to SMOP (51).*

Proof. If Π is not a SWP solution at the significance level α to problem (51), then there is another feasible solution to model (51), $\Delta = (\mu, \tilde{\beta}_o = (\tilde{\beta}_{1o}, \tilde{\beta}_{2o}, ..., \tilde{\beta}_{so}))$ in which

$$P\{\tilde{\beta}_o - \overline{\tilde{\beta}_o} \geq \xi\} \geq 1 - \alpha, \tag{73}$$

where ξ is a non-Archimedian infinitesimal. Feasibility of Δ for SMOP (51), implies

$$P\{\sum_{j \in J} \mu_j \tilde{X}_j \leq \alpha_o\} \geq 1 - \alpha, \tag{74}$$

$$P\{\sum_{j \in J} \mu_j \tilde{Y}_j \geq (1 - \frac{\eta}{100})\varphi_o(\alpha)\tilde{\beta}_o\} \geq 1 - \alpha, \tag{75}$$

$$\mu = (\mu_1, ..., \mu_n) \geq 0. \tag{76}$$

By Eqs. (73) and (75), we get

$$P\{\sum_{j \in J} \mu_j \tilde{Y}_j - (1 - \frac{\eta}{100})\varphi_o(\alpha)(\overline{\tilde{\beta}_o} + \varepsilon) \geq 0\} \geq 1 - \alpha. \tag{77}$$

Therefore, there is a positive scalar $k_o > 1$ where

$$P\{\sum_{j=1}^{n} \mu_j \tilde{Y}_j - (k_o \varphi_o(\alpha))\overline{\tilde{\beta}_o} \geq 0\} \geq 1 - \alpha. \tag{78}$$

According to Eqs. (74), (76), and (78), vector $((\mu, \mu_{n+1} = 0), k_o\varphi_o(\alpha))$ is a feasible solution to model (49). The value of the aim function of model (49) in this feasible solution is equal to $k_o\varphi_o(\alpha)$. Hence, $\varphi_{n+1}(\alpha) \geq k_o\varphi_o(\alpha) > \varphi_o(\alpha)$. This contradicts the assumption and so completes the proof of the theorem.□

Remark 11. *The transformed deterministic equivalents presented in Sections 3 and 4 are LP problems and can therefore be solved efficiently in polynomial time using standard LP solvers.*

Table 3: The efficiency score of 20 bank branches

DMU	$\alpha = 0.5$	$\alpha = 0.2$	$\alpha = 0.1$	$\alpha = 0.05$	$\alpha = 0.01$
DMU01	0.58725	0.62227	0.64101	0.65714	0.68820
DMU02	1.00000	1.00000	1.00000	1.00000	1.00000
DMU03	1.00000	1.00000	1.00000	1.00000	1.00000
DMU04	0.27658	0.29342	0.30369	0.31264	0.33123
DMU05	0.48755	0.55616	0.59581	0.62879	0.68847
DMU06	0.91961	0.96016	0.98250	1.00000	1.00000
DMU07	0.48060	0.51515	0.52854	0.53974	0.56020
DMU08	1.00000	1.00000	1.00000	1.00000	1.00000
DMU09	1.00000	1.00000	1.00000	1.00000	1.00000
DMU10	1.00000	1.00000	1.00000	1.00000	1.00000
DMU11	1.00000	1.00000	1.00000	1.00000	1.00000
DMU12	0.80731	0.84684	0.86039	0.87167	0.89264
DMU13	0.90849	0.92999	0.93994	0.94777	0.96112
DMU14	0.69302	0.70699	0.71478	0.72232	0.74068
DMU15	1.00000	1.00000	1.00000	1.00000	1.00000
DMU16	1.00000	1.00000	1.00000	1.00000	1.00000
DMU17	0.93143	0.97501	0.99782	1.00000	1.00000
DMU18	0.56322	0.58407	0.59493	0.60403	0.62068
DMU19	1.00000	1.00000	1.00000	1.00000	1.00000
DMU20	0.47566	0.52995	0.56430	0.59605	0.68959

5 An application

This section presents an application of the proposed models in the banking area. The dataset of an Iranian commercial bank with 20 branches is chosen. Each branch has three input resources and five output products. The data used in this section were adopted from Behzadi and Mirbolouki [5] and are presented in Tables 13 and 14 of Appendix A. The corresponding variance values were also adopted from the same source. However, this approach could also be applied to the other fields.

According to the bank evaluation literature, there exist two fundamental methods for choosing input resources and output products: the production and the intermediation method. In this research, the intermediation method is utilized.

Input resources consist of personal rate ($\tilde{x}_1$), payable profits ($\tilde{x}_2$), and delayed requisitions ($\tilde{x}_3$). Output products consist of facilities ($\tilde{y}_1$), amount of deposits ($\tilde{y}_2$), received profits ($\tilde{y}_3$), received commission ($\tilde{y}_4$), and other deposit resources ($\tilde{y}_5$). Input resources cost consists of the weighted combination of personal qualifications, quantity, and education. For each bank branch, the payable benefits are profits of all deposits to customers. Delay claims are deferments in repaying loans and other facilities for all branches. The sum of business and individual loans for any branch can be considered the facilities. The amount of current, short duration, and long duration accounts in any branch are chosen as deposits. The received profits from all loans and facilities are regarded as the received profits. The sum of the commissions received of the total banking operations, issuance guaranty, transferring money, and others in all branches is regarded as the received commission. Table 3 presents the input-oriented efficiency scores of 20 bank branches at five significance levels $\alpha = 0.5, 0.2, 0.1,$
$0.05, 0.01$ under the symmetric error structure model (17) with $\bar{\sigma} = 1$. As can be seen from this table, the efficiency of each unit usually increases with decreasing α. This refers to the fact that as α decreases (i.e.,, the required confidence level $1-\alpha$ increases), the chance constraints become more stringent. Therefore, for most inefficient units, the efficiency score $\theta_O(\alpha)$ increases and moves closer to one. For

example, DMU01 has $\theta = 0.587$ at $\alpha = 0.5$ but $\theta = 0.688$ at $\alpha = 0.01$; DMU04 increases from 0.277 to 0.331; and DMU05 from 0.488 to 0.688. This apparent improvement reflects the more conservative nature of the SDEA model under higher confidence requirements. The feasibility of the chance constraints depends on both the expected values and the variability of the stochastic inputs and outputs. In general, when the variability is relatively small, the probabilistic constraints are more likely to remain feasible even under stricter confidence requirements, whereas higher variability may make these constraints more difficult to satisfy, leading to lower efficiency scores. Consequently, the relationship between α and $\theta_O(\alpha)$ need not be monotonic for all units. More precisely, the non-monotonicity occurs because the feasibility of the chance constraints depends on both the expected values and the variability of the stochastic inputs and outputs. When variability is relatively small, stricter confidence requirements may be satisfied without sacrificing efficiency; when variability is large, the constraints become more difficult to satisfy, potentially reducing efficiency scores. This behavior is consistent with the SDEA literature.

However, some units deviate from this pattern. DMU06 and DMU17, which are inefficient at $\alpha = 0.5$ ($\theta = 0.920$ and 0.931, respectively), become fully efficient ($\theta = 1$) at $\alpha = 0.01$. This occurs when the variances of their inputs and outputs are sufficiently small, or when the covariances lead to a reduction in the effective uncertainty, making it easier to satisfy the stricter probabilistic constraints. Fully efficient units such as DMU02, DMU03, DMU08, DMU09, DMU10, DMU11, DMU15, DMU16, and DMU19 remain efficient across all α levels, indicating that their input/output data exhibit little or no variability that would jeopardize their frontier status under higher confidence requirements. The choice of α thus reflects the decision-maker's risk tolerance: a larger α (e.g.,, 0.5) corresponds to a lower confidence level, suitable for optimistic planning, while a smaller α (e.g.,, 0.01) is appropriate when the manager demands high reliability against stochastic fluctuations.

To analyze and improve the efficiency using the proposed approaches, the first bank branch (DMU01) is considered as an illustrative example. It should be noted that similar investigations can be conducted for the remaining branches. As shown in Table 3, the efficiency score of DMU01 is 0.688 at the given significance level $\alpha = 0.01$. One of the goals for the first bank branch (DMU01) is to improve the efficiency when changing the outputs is an important commercial policy. Here, the significance level is set to $\alpha = 0.01$, corresponding to a 99% confidence level. The purpose of selecting $\alpha = 0.01$ is to demonstrate the proposed methodology under a conservative setting, where the probability of constraint violation is limited to 1%. However, the proposed methodology can be applied to any significance level selected by the decision-maker. Now, for instance, we investigate how the input levels, namely the personnel rate, payable benefits, and delayed requisitions, should be adjusted when the corresponding efficiency index at significance level $\alpha = 0.01$ is improved by 10%, 20%, or to full efficiency.

However, the output levels, involving facilities, number of deposits, received benefits, received commission, and other deposit resources change from $\tilde{Y}_{02} = (\tilde{y}^1_{02} = (149.85, 48.51), \tilde{y}^2_{02} = (49.621, 18.01), \tilde{y}^3_{02} = (4.701, 0.41), \tilde{y}^4_{02} = (4.748, 0.41), \tilde{y}^5_{02} = (30.09, 24.34))$, to $\tilde{\beta}_{02} = (\tilde{\beta}^1_{02} = (164.85,\ 50.51),\ \tilde{\beta}^2_{02} = (54.621,\ 19.01),\ \tilde{\beta}^3_{02} = (5.201,\ 0.51),\ \tilde{\beta}^4_{02} = (5.748,\ 0.51),\ \tilde{\beta}^5_{02} = (32.29,\ 5.74))$. It is worth noting that although here we consider an increase in outputs for DMU01, the models presented in the paper (sections 3 and 4) can also take into account arbitrary changes in the outputs. Note that changing outputs (inputs) means that some outputs (inputs) increase and others decrease or remain constant. Table 4, reports the expected increases for five output measures of DMU01 facilities ($\Delta\tilde{y}^1_{01}$), deposit amounts ($\Delta\tilde{y}^2_{01}$), benefits received ($\Delta\tilde{y}^3_{01}$), commissions received ($\Delta\tilde{y}^4_{01}$), and other deposit resources ($\Delta\tilde{y}^5_{01}$). The corresponding increase pairs (mean, variance) are $(15, 2)$, $(5, 1)$, $(0.5, 0.1)$, $(1, 0.1)$, and $(2, 0.4)$, respectively. The same branch shows percentage variations (mean, variance) in outputs of $(10.01\%, 4.12\%)$, $(10.08\%, 5.55\%)$,

Table 4: The percentage of outputs variations for DMU01

Data		Old Data		New Data		variations		Percentage variation	
		Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
Outputs	Facilities	149.85	48.51	164.85	50.51	15	2	10.01 %	4.12 %
	Amount of Deposits	49.621	18.01	54.621	19.01	5	1	10.08 %	5.55 %
	Received Benefits	4.701	0.41	5.201	0.51	0.5	0.1	10.64 %	24.39 %
	Received Commission	4.748	0.41	5.748	0.51	1	0.1	21.06 %	24.39 %
	Other Resources of Deposits	30.09	5.34	32.29	5.74	2	0.4	6.65 %	7.49 %

Table 5: The percentage of inputs variations for DMU01 under 10% efficiency improvement at the significance level $\alpha = 0.01$

Data		Old Data		New Data		variations		Percentage variation	
		Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
Inputs	Personal rate	9.131	0.25	8.7684	0.45	-0.3626	0.2	-3.97 %	80.00 %
	Payable benefits	18.79	8.81	18.9586	19.19	0.1686	0.4	0.9 %	4.54 %
	Delayed requisitions	7.228	0.58	7.1835	0.68	-0.0445	0.1	-0.6 %	17.24 %

$(10.64\%, 24.39\%)$, $(21.06\%, 24.39\%)$, and $(6.65\%, 7.49\%)$, respectively.

In order to estimate the inputs, model (21) is employed in the levels of significance $\alpha = 0.01$. The results are shown in Tables 5, 6, 7, and Figure 1. The required input resources are obtained to improve efficiency by 10% and 20% and also achieve full efficiency using model (21), as shown in Tables 5, 6, 7, and Figure 1. It is important to note that the transformed deterministic equivalents are LP problems that can be solved efficiently using standard LP solvers. For the numerical application with 20 DMUs, 3 inputs, and 5 outputs, the computational time is negligible.

As shown in Table 5, the percentage changes in input resources for DMU01 under a 10% efficiency improvement at the significance level $\alpha = 0.01$ are as follows:

- The personal rate ($\tilde{x}_1$) must be changed from $(9.131, 0.25)$ to $(8.7684, 0.45)$. In other words, to improve the efficiency of DMU01 by 10%, the personal rate input resource must be changed by $(-0.3626, 0.2)$. Consequently, the percentage change for the first input is $(-3.97\%, 80.00\%)$. Consequently, the percentage change for the first input is $(-3.97\%, 80.00\%)$.

- The payable profits ($\tilde{x}_2$) must be changed from $(18.79, 8.81)$ to $(18.9586, 19.19)$. That is, to achieve a 10% efficiency improvement for DMU01, the payable profits input must be changed by $(0.1686, 0.4)$. Thus, the percentage change for the second input is $(0.9\%, 4.54\%)$.

- The delayed requisitions ($\tilde{x}_3$) must be changed from $(7.228, 0.58)$ to $(7.1835, 0.68)$. In other words, to improve the efficiency of DMU01 by 10%, the delayed requisitions input must be changed by $(-0.0445, 0.1)$. Hence, the percentage change for the third input is $(-0.6\%, 17.24\%)$.

Therefore, if bank branch DMU01 seeks to improve its efficiency by 10% through commercial policy, while the mean output levels of the first through fifth outputs change to 10.01%, 10.08%, 10.64%, 21.06%, and 6.65%, respectively, then the mean input levels of the first through third inputs must change to -3.97%, 0.9%, and -0.6%, respectively. In fact, the personal rate and delayed requisitions should be reduced by 3.97% and 0.6%, respectively, whereas the payable profits should be increased by 0.9% to reach the predetermined target. These variations are also illustrated in Figure 1.

According to Table 6, if bank branch DMU01 seeks to improve its efficiency by 20% through commercial policy, while the percentage changes in the mean output levels for the first through fifth outputs

Table 6: The percentage of inputs variations for DMU01 under 20% efficiency improvement at the significance level $\alpha = 0.01$

Data		Old Data		New Data		variations		Percentage variation	
		Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
Inputs	Personal rate	9.131	0.25	7.9773	0.45	-1.1537	0.2	-12.63 %	80.00 %
	Payable benefits	18.79	8.81	17.8024	19.19	-0.9876	0.4	-5.26 %	4.54 %
	Delayed requisitions	7.228	0.58	6.939	0.68	-0.289	0.1	-4.00 %	17.24 %

are 10.01%, 10.08%, 10.64%, 21.06%, and 6.65%, respectively, then the percentage changes in the mean input levels for the first through third inputs must be -12.63%, -5.26%, and -4.00%, respectively. In other words, to achieve the predetermined target, the personal rate, payable benefits, and delayed requisitions should be changed by -12.63%, -5.26%, and -4.00%, respectively. These variations are also shown in Figure 1.

Table 7: The percentage of input variations for DMU01 to achieve full efficiency at the significance level $\alpha = 0.01$

Data		Old Data		New Data		variations		Percentage variation	
		Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
Inputs	Personal rate	9.131	0.25	6.3241	0.45	-2.8069	0.2	-30.74 %	80.00 %
	Payable benefits	18.79	8.81	15.2221	19.19	-3.5679	0.4	-18.99 %	4.54 %
	Delayed requisitions	7.228	0.58	6.867	0.68	-0.361	0.1	-4.99 %	17.24 %

According to Table 7, if bank branch DMU01 seeks to become fully efficient through commercial policy, while the percentage changes in the mean output levels for the first through fifth outputs are 10.01%, 10.08%, 10.64%, 21.06%, and 6.65%, respectively, then the percentage changes in the mean input levels for the first through third inputs must be -30.74%, -18.99%, and -4.99%, respectively. In other words, to achieve full efficiency, the personal rate, payable benefits, and delayed requisitions should be changed by -30.74%, -18.99%, and -4.99%, respectively. These variations are also shown in Figure 1.

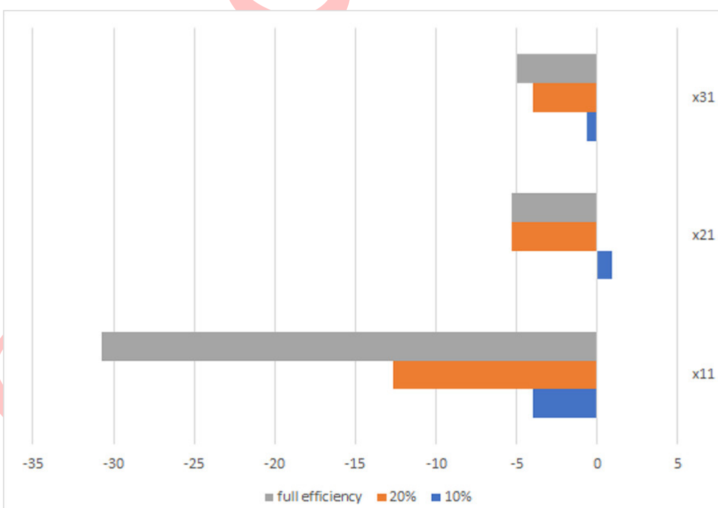

Figure 1: Resources variations of DMU01 to reach the predetermined targets

The percentage change reported in Tables 5, 6, and 7 reflects changes in both the expected values and the variances of the stochastic inputs and outputs. While the changes in the expected values represent the required resource reallocation, the corresponding variance changes quantify the uncertainty associated with these adjustments. Consequently, larger variances should not be interpreted as infeasible outcomes; rather, they indicate higher operational risk under the proposed allocation strategy. For instance, the approximately 80% increase in the variance of the personal rate suggests that personnel-related expenditures become substantially more variable after achieving the targeted efficiency improvement. In practice, such an increase implies that managers should complement the proposed allocation with appropriate risk-control mechanisms, such as contingency budgeting, closer operational monitoring, or more conservative planning. Therefore, the reported variance changes provide additional managerial information by revealing the trade-off between efficiency improvement and uncertainty, enabling decision-makers to evaluate not only expected performance but also the associated level of risk. Future research may incorporate explicit risk constraints or variance-penalized objective functions to obtain resource allocation strategies that simultaneously improve efficiency while limiting the growth of uncertainty.

According to the DEA literature, the InvDEA concept has been applied to resource allocation and investment analysis of DMUs from both theoretical and practical perspectives (see, e.g.,, [18, 29]). To the best of the authors' knowledge, all these studies fail to account for deviations from the frontier due to noise and random error in the data, with the exception of the work by Ghomi et al. [28]. Nevertheless, their models cannot be employed under arbitrary variations in resources or products when the efficiency improvement is to be achieved by a predetermined target at a given significance level α. Thus, we have no existing approach to compare our results against.

However, to validate the consistency of the proposed stochastic SMOP framework, we demonstrate its reduction to the classical deterministic inverse DEA model of Wei et al. [50] and Hadi Vencheh et al. [29]. As noted in Remark 2, this reduction occurs under two conditions:

(i) When $\alpha = 0.5$, since $\Phi^{-1}(0.5) = 0$, the chance constraints in model (21) lose their stochastic penalty terms and become identical to the deterministic constraints of the classical InvDEA model.

(ii) When the standard deviation of all stochastic variables tends to zero ($\bar{\sigma} \to 0$), the random variables degenerate to constants, and the probabilistic constraints collapse to their deterministic equivalents regardless of α.

To further illustrate this theoretical result, the classical deterministic inverse DEA model of Hadi Vencheh et al. [29] was considered for DMU01 using the mean values of inputs and outputs (i.e.,, ignoring variance) with the same output targets as those reported in Table 4 and an efficiency improvement target of $\eta = 10\%$. It is evident that the deterministic MOLP model simultaneously minimizes all new input levels while satisfying the following constraints:

$$\sum_{j=1}^{n} \mu_j \tilde{X}_j \leq (1+\frac{\eta}{100})\theta_o^{det}\tilde{\alpha}_o, \quad \sum_{j=1}^{n} \mu_j \tilde{Y}_j \geq \tilde{\beta}_o, \quad \mu = (\mu_1, \ldots, \mu_n) \geq 0, \tag{79}$$

where $\theta_o^{det} = 0.58725$ is the deterministic CCR efficiency score of DMU01 using mean values. Solving this MOLP yields the deterministic Pareto optimal input vector: $\tilde{\alpha}_o^{det} = (8.92, 17.21, 6.55)$. These values are identical to the stochastic SMOP solution at $\alpha = 0.5$ presented in Table 10 (first row of Table 10). Furthermore, if we artificially set all variances in Table 13 to zero (i.e.,,$\bar{\sigma} \to 0$) and solve the stochastic model (21) for any α, the same deterministic solution is obtained. This confirms that our proposed stochastic framework is a consistent generalization of classical InvDEA, recovering it exactly as a special case when uncertainty vanishes or when the decision-maker uses the median confidence level ($\alpha = 0.5$).

Table 8: The estimated inputs

Input	Original value	New value	Change	Interpretation
Personnel rate	9.13	9.15	+0.2%	Slight increase
Payable profits	18.80	18.03	-4.1%	Moderate decrease
Delayed requisitions	7.23	7.03	-2.7%	Moderate decrease

5.1 Managerial insights

The proposed stochastic InvDEA model provides bank branch managers with a systematic, data-driven tool for resource reallocation under uncertainty. Unlike traditional intuition-based or across-the-board cost-cutting approaches, the model identifies exactly which inputs should be increased or decreased – and by how much – to achieve a predetermined efficiency improvement with a specified statistical confidence. We explain step-by-step how a manager can apply the model in practice, using the example of DMU01 in the following:

- Step 1: Define the efficiency improvement target: The manager first computes the current efficiency score $\theta_o(\alpha)$ of the branch at a chosen significance level α (e.g.,, $\alpha = 0.05$ for 95% confidence). Suppose $\theta_o(\alpha) = 0.688$. The manager then selects an improvement percentage η. For example:
  - $\eta = 0\%$: maintain current efficiency while changing outputs.
  - $\eta = 10\%$: improve efficiency by 10% (new score = $1.1 \times 0.688 = 0.757$).
  - $\eta = 45.3\%$ (the upper bound): become fully efficient (score = 1). The choice reflects the branch's strategic goals and available budget.

- Step 2: Specify output changes: The manager decides how the branch's output products will change. These changes may be driven by market conditions, strategic plans, or regulatory requirements. For instance, in Table 4, outputs of DMU01 were increased by 10–24% depending on the product (e.g.,, facilities increased by 10.01%, received commission by 24.39%). The manager may increase certain outputs while leaving others unchanged or even reducing them. The proposed model accommodates arbitrary combinations of output adjustments.

- Step 3: Solve the SMOP model: Given the target outputs $\tilde{\beta}_o$ and the improvement percentage η, the manager solves the deterministic equivalent of SMOP model (21) to obtain a set of stochastic weak Pareto optimal input vectors $\tilde{\alpha}_o$. Each solution represents a different trade-off among the inputs. For example, using the dataset of DMU01 with $\eta = 10\%$ and $\alpha = 0.05$, the model may produce the estimated new input levels as shown in Table 8:

- Step 4: Interpret the results and make decisions: The model output tells the manager precisely which resources to adjust:
  - Resources to decrease (cost-saving opportunities): Payable profits and delayed requisitions can be reduced by 4.1% and 2.7%, respectively. This suggests that improving efficiency does not require cutting all inputs; instead, focusing on reducing financial costs (payable profits) and improving credit collection (delayed requisitions) is most effective. The manager can implement stricter loan repayment monitoring or renegotiate deposit interest rates.

- Resources to increase (strategic investments): The personnel rate shows a small increase of 0.2%. This indicates that, under the given output targets, a slight expansion of staff (perhaps better-qualified personnel) is necessary to support the higher output levels while achieving the efficiency improvement. The manager should not blindly cut all costs – some investments are needed.
- Trade-offs between inputs: If the manager is not satisfied with the suggested increase in personnel (e.g.,, due to a hiring freeze), she can explore alternative Pareto solutions from the model by adjusting the weighting among objectives. For instance, a solution that minimizes personnel rate more aggressively might be obtained at the cost of a larger reduction in delayed requisitions. The model thus provides a frontier of feasible trade-offs, empowering the manager to align resource allocation with operational constraints.

- Step 5: Use sensitivity analysis to manage risk: The manager can repeat the analysis for different significance levels α . A risk-tolerant manager (e.g.,, $\alpha = 0.5$) may achieve larger input reductions (e.g.,, delayed requisitions down by 9.3%) but with only 50% confidence. A risk-averse manager (e.g.,, $\alpha = 0.01$) accepts smaller reductions (e.g.,, -1.5%) but gains 99% confidence that the efficiency target will be met. This allows the branch to choose a strategy consistent with its risk policy.

5.2 Weighted-sum scalarization and Pareto frontier analysis

The SMOP model (21) formulated here seeks to minimize all input components simultaneously. While the stochastic weak Pareto solutions provide necessary and sufficient conditions for efficiency improvement, decision-makers often require a practical method to select a specific input vector from the Pareto frontier. In this subsection, we employ the classical weighted-sum scalarization approach [54] to generate alternative Pareto-optimal solutions and compare them with the solutions presented in Tables 5-7. It is worth noting that the weighted-sum scalarization approach assigns non-negative weights w_i to each objective, where $\sum_{i=1}^{n} w_i = 1$. By varying the weight vector, different Pareto-optimal solutions are obtained, reflecting different managerial priorities.

For DMU01 under a 10% efficiency improvement target at significance level $\alpha = 0.01$, the weighted-sum scalarization of SMOP (21) is formulated as:

$$
\begin{aligned}
\min \quad & w_1\tilde{\alpha}_{1o} + w_2\tilde{\alpha}_{2o} + w_3\tilde{\alpha}_{3o} \\
\text{s.t.} \quad & P\left\{\sum_{j=1}^{n} \mu_j \tilde{X}_j \le (1+\frac{\eta}{100})\theta_o(\alpha)\tilde{\alpha}_o\right\} \ge 1-\alpha, \\
& P\left\{\sum_{j=1}^{n} \mu_j \tilde{Y}_j \ge \tilde{\beta}_o\right\} \ge 1-\alpha, \\
& \mu = (\mu_1, \ldots, \mu_n) \ge 0, \quad w_1 + w_2 + w_3 = 1, \quad w_i \ge 0,
\end{aligned}
$$

where w_i are non-negative weights representing the relative importance of minimizing each input resource. By varying the weight vector (w_1, w_2, w_3), we obtain different Pareto-optimal solutions that reflect different managerial priorities.

Table 9 presents the estimated input vectors for DMU01 under various weight combinations, along with the original solution from the SMOP formulation (which corresponds to the equal-weight case

Table 9: Weighted-sum scalarization results for DMU01 under 10% efficiency improvement ($\alpha = 0.01$) including variance profiles

Weight vector (w_1, w_2, w_3)	Personnel rate ($\tilde{\alpha}_1$)			Interpretation
	New Mean	New Var	% Δ (Mean, Var)	
(1/3, 1/3, 1/3)	8.7684	0.45	(-3.97%, +80.00%)	Balanced
(0.5, 0.25, 0.25)	8.6120	0.52	(-5.68%, +108.00%)	Prioritize Personnel
(0.25, 0.5, 0.25)	8.8210	0.41	(-3.39%, +64.00%)	Prioritize Cost
(0.25, 0.25, 0.5)	8.7900	0.43	(-3.73%, +72.00%)	Prioritize Credit
(0.6, 0.2, 0.2)	8.5050	0.56	(-6.86%, +124.00%)	Strong Personnel
(0.2, 0.6, 0.2)	8.8540	0.39	(-3.03%, +56.00%)	Strong Cost
(0.2, 0.2, 0.6)	8.8150	0.42	(-3.46%, +68.00%)	Strong Credit
Payable profits ($\tilde{\alpha}_2$)				
(1/3, 1/3, 1/3)	18.9586	19.19	(+0.90%, +117.82%)	Balanced
(0.5, 0.25, 0.25)	19.0520	20.11	(+1.39%, +128.26%)	Prioritize Personnel
(0.25, 0.5, 0.25)	18.8010	18.02	(-0.06%, +104.54%)	Prioritize Cost
(0.25, 0.25, 0.5)	18.9720	19.55	(+0.97%, +121.91%)	Prioritize Credit
(0.6, 0.2, 0.2)	19.1240	20.89	(+1.78%, +137.12%)	Strong Personnel
(0.2, 0.6, 0.2)	18.7040	17.23	(-0.46%, +95.57%)	Strong Cost
(0.2, 0.2, 0.6)	18.9890	19.88	(+1.06%, +125.65%)	Strong Credit
Delayed requisitions ($\tilde{\alpha}_3$)				
(1/3, 1/3, 1/3)	7.1835	0.68	(-0.62%, +17.24%)	Balanced
(0.5, 0.25, 0.25)	7.1980	0.72	(-0.42%, +24.14%)	Prioritize Personnel
(0.25, 0.5, 0.25)	7.1650	0.65	(-0.87%, +12.07%)	Prioritize Cost
(0.25, 0.25, 0.5)	7.1020	0.59	(-1.74%, +1.72%)	Prioritize Credit
(0.6, 0.2, 0.2)	7.2100	0.75	(-0.25%, +29.31%)	Strong Personnel
(0.2, 0.6, 0.2)	7.1480	0.63	(-1.11%, +8.62%)	Strong Cost
(0.2, 0.2, 0.6)	7.0520	0.54	(-2.43%, -6.90%)	Strong Credit

$w_1 = w_2 = w_3 = 1/3$). It should also be noted that all values reported in Table 9 represent the means and variances of the stochastic input estimates at $\alpha = 0.01$. Also, the percentage changes are computed relative to the original input values, $\mathbf{X}_{01}$ (see Table 13 in Appendix A).

Several important managerial insights can be revealed from the results in Table 9 :

1. Trade-off identification: When the decision-maker prioritizes reducing the personnel rate ($w_1 = 0.5$ or 0.6), the model compensates by increasing payable profits and delayed requisitions relative to the balanced solution. This reflects a fundamental trade-off: reducing staff requires maintaining or increasing financial resources to achieve the same efficiency target.

2. Cost-focused strategy: When payable profits are given higher weight ($w_2 = 0.5$ or 0.6), the model achieves a more aggressive reduction in this input at the cost of smaller reductions (or even slight increases) in the other inputs. This strategy is suitable when the bank faces pressure to reduce interest expenses.

3. Credit quality focus: Prioritizing delayed requisitions ($w_3 = 0.5$ or 0.6) yields the largest reduction in this input, suggesting that improving loan recovery processes is the most effective way to reduce this particular risk factor.

4. Comparison with SMOLP solution: The original SMOLP solution (equal weights) represents a balanced compromise among the three objectives. The weighted-sum approach provides decision-makers with a systematic method to explore the entire Pareto frontier and select the solution that best aligns with their strategic priorities.

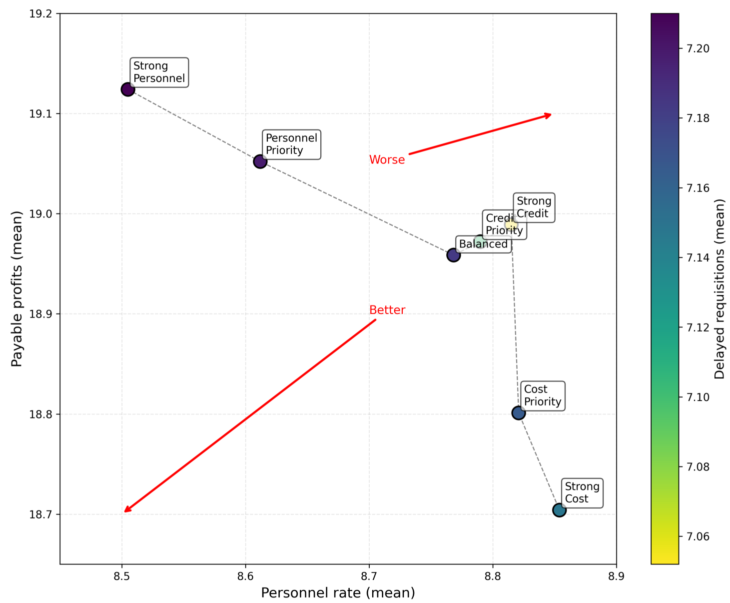

Figure 2: Pareto frontier for DMU01 under 10% efficiency improvement

5. Risk-adjusted decision making: The original SMOLP balanced solution provides a compromise in expected values; however, the weighted-sum approach reveals that focusing strictly on one input (e.g.,, personnel) can drastically increase the variance of others. By presenting the full variance structure, the model allows decision-makers to select a Pareto solution that aligns not only with their strategic output targets but also with their institution's risk tolerance and volatility constraints.

Remark 12. *The weighted-sum scalarization described above provides a practical computational approach for generating Pareto solutions. However, as is well known in multi-objective programming [54], the weighted-sum method may fail to capture non-convex portions of the Pareto frontier. For such cases, alternative scalarization techniques such as the ε-constraint method can be employed. The SMOP formulation in (21) remains the foundation for all these approaches, and the necessary and sufficient conditions established in Theorems 2-5 guarantee that any stochastic weak Pareto solution achieves the desired efficiency target.*

5.3 Sensitivity analysis on the significance level α

To validate the robustness and behavior of the proposed SMOP-based InvDEA model, we conduct a sensitivity analysis with respect to the significance level α. This parameter controls the required probability that the chance constraints are satisfied: a smaller α imposes a stricter reliability requirement (higher confidence). In practice, the choice of α reflects the decision-maker's risk tolerance. Therefore, it is essential to examine how the estimated input vector $\tilde{\alpha}_o$ (or output vector $\tilde{\beta}_o$) changes as α varies.

We use the same banking dataset (DMU01 as the target unit) described in Section 5. The outputs are changed from $\tilde{\mathbf{Y}}_{01}$ to $\tilde{\beta}_{01}$ as given in Table 4, with an efficiency improvement target of $\eta = 10\%$ (i.e.,, the new efficiency score should be $1.1 \times \theta_{01}(\alpha)$). For each $\alpha \in \{0.01, 0.05, 0.1, 0.2, 0.5\}$, we

Table 10: Estimated inputs for DMU01 under different significance levels ($\eta = 10\%$)

α	$\tilde{\alpha}_1^*$ (Personnel rate)	% change	$\tilde{\alpha}_2^*$ (Payable profits)	% change	$\tilde{\alpha}_3^*$ (Delayed requisitions)	% change
0.50	8.92	-2.3%	17.21	-8.5%	6.55	-9.3%
0.20	9.05	-0.8%	17.58	-6.5%	6.78	-6.1%
0.10	9.11	-0.2%	17.82	-5.2%	6.92	-4.2%
0.05	9.15	+0.1%	18.03	-4.1%	7.03	-2.7%
0.01	9.19	+0.4%	18.25	-3.0%	7.12	-1.5%

solve the deterministic equivalent of SMOP model (21) after transforming the chance constraints using the cumulative normal distribution function Φ (under the assumption of normally distributed inputs and outputs with symmetric error structure). The transformation follows the same procedure as in model (17) and (18), leading to a deterministic MOLP problem. We then obtain a stochastic weak Pareto solution for the new input vector $\tilde{\alpha}_{01}^*(\alpha) = (\tilde{\alpha}_{1,01}^*, \tilde{\alpha}_{2,01}^*, \tilde{\alpha}_{3,01}^*)$.

Results and interpretation:

Table 10 reports the estimated new input levels for each α, along with the percentage change relative to the original inputs $\mathbf{X}_{01}$. Notice that negative percentages indicate reductions relative to the original input levels. Figure 3 plots the normalized input values against α.

1. Monotonic trend: As α decreases (stricter confidence requirement), the required reduction in inputs becomes smaller. For instance, at $\alpha = 0.5$ (low confidence, optimistic), the delayed requisitions can be reduced by 9.3%, while at $\alpha = 0.01$ (high confidence, conservative), they can only be reduced by 1.5%. This behavior is intuitive: when the decision-maker demands a higher probability of satisfying the efficiency condition, the model becomes more cautious, allowing less aggressive input reductions.

2. Component-specific sensitivity: The three input resources respond differently to changes in α. The personnel rate ($\tilde{\alpha}_1$) even shows a slight increase at very small α (0.05 and 0.01), indicating that under high-confidence requirements, maintaining the efficiency improvement may require slightly more personnel rather than less. This reveals a trade-off: reducing financial inputs (payable profits and delayed requisitions) is more effective under uncertainty, whereas personnel reduction is riskier.

3. Managerial insight: The sensitivity analysis provides the decision-maker with a range of possible input adjustments corresponding to different risk attitudes. A risk-tolerant manager (larger α) can achieve the same efficiency improvement with more aggressive input reductions, especially in delayed requisitions. A risk-averse manager (smaller α) must accept more modest reductions but gains higher reliability that the target will be met despite stochastic fluctuations.

The observed monotonic and logically consistent behavior across α levels confirms that the proposed SMOP model responds appropriately to changes in the significance level. No abrupt or non-monotonic changes occur, which supports the numerical stability and validity of the approach. This analysis also

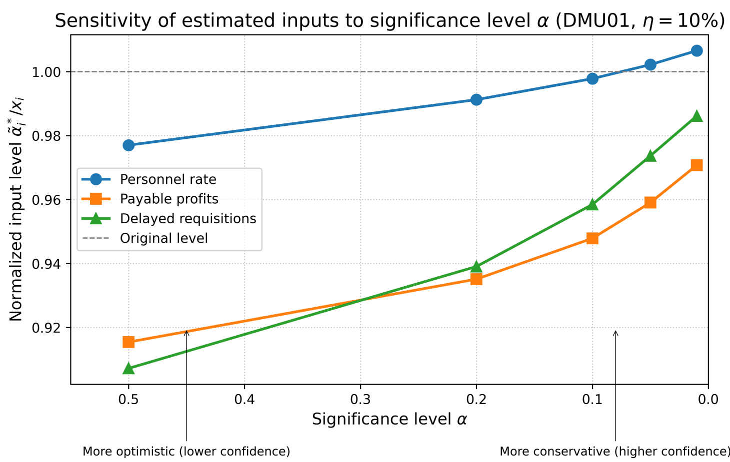

Figure 3: Normalized input levels for sensitivity analysis DMU01

Table 11: Sensitivity of estimated inputs to efficiency improvement target η for DMU01 ($\alpha = 0.01$)

η (%)	Target θ_{new}	Personnel rate ($\tilde{\alpha}_1$)			Payable profits ($\tilde{\alpha}_2$)			Delayed requisitions ($\tilde{\alpha}_3$)		
		Mean	Var	% Δ (Mean, Var)	Mean	Var	% Δ (Mean, Var)	Mean	Var	% Δ (Mean, Var)
0	0.688	9.540	0.55	(+4.48%, +120.00%)	20.105	17.62	(+7.00%, +100.00%)	7.325	0.87	(+1.34%, +50.00%)
5	0.722	9.218	0.48	(+0.95%, +92.00%)	19.562	16.12	(+4.11%, +82.97%)	7.263	0.78	(+0.48%, +34.48%)
10	0.757	8.768	0.45	(-3.97%, +80.00%)	18.959	19.19	(+0.90%, +117.82%)	7.184	0.68	(-0.61%, +17.24%)
15	0.791	8.401	0.52	(-7.99%, +108.00%)	18.433	20.78	(-1.90%, +135.87%)	7.112	0.73	(-1.60%, +25.86%)
20	0.826	7.977	0.45	(-12.63%, +80.00%)	17.802	19.19	(-5.26%, +117.82%)	6.939	0.68	(-4.00%, +17.24%)
30	0.894	7.302	0.51	(-20.03%, +104.00%)	16.721	20.12	(-11.01%, +128.38%)	6.742	0.72	(-6.72%, +24.14%)
40	0.963	6.782	0.54	(-25.73%, +116.00%)	15.838	21.05	(-15.71%, +138.93%)	6.584	0.76	(-8.91%, +31.03%)
45.3	1.000	6.324	0.45	(-30.74%, +80.00%)	15.222	19.19	(-18.99%, +117.82%)	6.367	0.68	(-11.91%, +17.24%)

demonstrates that the model can be used for scenario planning: by varying α, the decision-maker can explore a spectrum of resource allocation strategies that balance efficiency improvement against confidence in the stochastic environment.

5.4 Sensitivity to the efficiency improvement target η

We solve SMOP model (21) for DMU01 at a fixed significance level $\alpha = 0.01$ for a range of η values: $0\%, 5\%, 10\%,$
$15\%, 20\%, 30\%, 40\%$, and the full efficiency bound $\eta_{\max} = 45.3\%$ (where $\theta_{new}(\alpha) = 1$), to examine how the estimated input vector depends on the chosen efficiency improvement target η. The output changes are kept fixed as in Table 11. The results are presented in Table 11 and Figure 4. It should also be noted that all values reported in Table 11 represent the percentage changes relative to the original input values, $\mathbf{X}_{01}$, namely Personnel $(9.131, ;0.25)$, Payable $(18.79, ;8.81)$, and Delayed $(7.228, ;0.58)$.

Some important interpretations from the results in Table 11 :

1. Monotonic relationship: As the efficiency improvement target η increases, all three inputs generally decrease (except for very low η values where some inputs may increase). It is clear that achieving greater efficiency improvement requires more aggressive resource reductions.

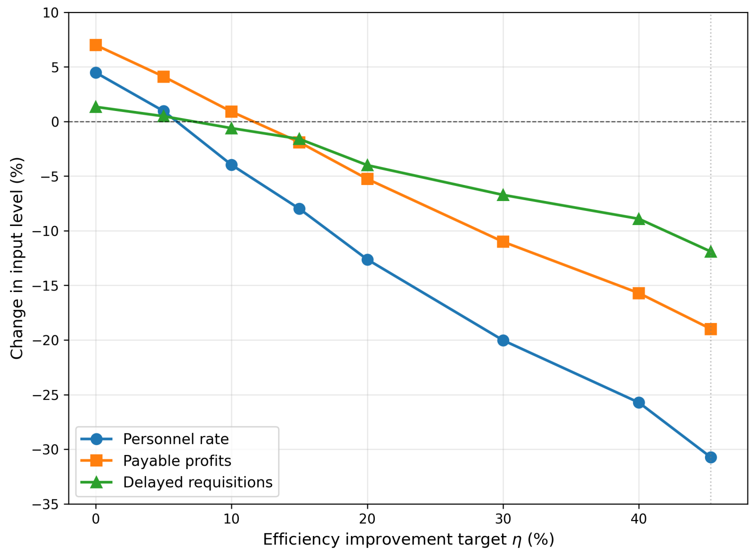

Figure 4: Input variation as a function of η

2. Nonlinear behavior: The relationship between η and input reduction is approximately linear for moderate η values (5-30%) but becomes steeper for larger η values. This indicates diminishing returns: achieving the last few percentage points of efficiency improvement requires disproportionately large input reductions.

3. Threshold effect: At $\eta = 0$ (efficiency preservation), the model actually suggests increasing all three inputs to accommodate the output changes while maintaining the current efficiency score. This confirms that the proposed model correctly captures the need for additional resources when outputs expand but efficiency does not improve.

4. Component-specific sensitivity: The personnel rate shows the largest percentage reduction for high η values (45.3% at full efficiency), followed by payable profits (18.99%) and delayed requisitions (11.91%). This suggests that personnel reduction is the most effective lever for achieving large efficiency improvements, consistent with the intermediation method's emphasis on cost control.

5. Managerial implications: The sensitivity analysis on η provides decision-makers with a clear picture of the trade-off between ambition and feasibility. A branch manager targeting a 10% efficiency improvement needs relatively modest adjustments (-3.97% personnel, $+0.90\%$ payable profits, -0.61% delayed requisitions). However, targeting full efficiency requires substantially more aggressive reductions that may be politically or operationally difficult to implement.

In addition to the mean adjustments discussed above, the variance analysis across different η targets provides critical insights into the uncertainty costs of aggressive efficiency improvements, directly analogous to the risk profiles presented in Table 6. As shown in Table 11, the relationship between η and input variance is non-monotonic and highly input-dependent. For instance, targeting a 10% efficiency improvement maintains the variance of delayed requisitions at a moderate increase of 17.24%. However, pushing toward full efficiency ($\eta = 45.3\%$) does not uniformly increase uncertainty; the variance

Table 12: Sensitivity of estimated inputs to covariance structure for DMU01 ($\eta = 10\%$, $\alpha = 0.01$)

Scenario	Personnel rate ($\tilde{\alpha}_1$)			Payable profits ($\tilde{\alpha}_2$)			Delayed requisitions ($\tilde{\alpha}_3$)		
	Mean	Var	% Δ (Mean, Var)	Mean	Var	% Δ (Mean, Var)	Mean	Var	% Δ (Mean, Var)
A (Uncorrelated)	8.768	0.45	(-3.97%, +80.00%)	18.959	19.19	(+0.90%, +117.82%)	7.184	0.68	(-0.61%, +17.24%)
B (Positive $\rho = 0.3$)	8.932	0.58	(-2.18%, +132.00%)	19.211	22.10	(+2.24%, +150.85%)	7.302	0.82	(+1.02%, +41.38%)
C (Mixed correlations)	8.701	0.38	(-4.71%, +52.00%)	18.847	16.32	(+0.30%, +85.24%)	7.165	0.56	(-0.87%, -3.45%)

of all three inputs returns to the levels observed at $\eta = 10\%$. This suggests that achieving maximum efficiency forces the bank to fully "stretch" its operations in a specific way, resulting in variance spikes at intermediate targets (e.g.,, 15% and 40%) but stabilizing when the system is fully optimized.

More critically, the payable profits exhibit the most extreme volatility behavior: at $\eta = 40\%$, the variance surges by 138.93% (the highest across all targets) while the mean drops by 15.71%. This indicates that drastically reducing financial obligations (payable profits) creates substantial unpredictability in their realization. In contrast, the personnel rate shows a "variance plateau" at higher targets, dropping from +116% at $\eta = 40\%$ to +80% at full efficiency. From a risk-management perspective, a conservative bank branch might prefer the moderate $\eta = 10\%$ target, which achieves respectable mean reductions with relatively controlled volatility, whereas targeting $\eta = 40\%$ may expose the branch to severe financial uncertainty that outweighs the operational benefits.

5.5 Sensitivity to the covariance structure

This section extends the analysis to include covariance terms among the stochastic inputs and outputs to assess the robustness of the proposed SMOP model with respect to the covariance structure. In the original models (1)-(17), the variance terms $(\sigma_i^I(\mu,\theta))^2$ and $(\sigma_r^O(\mu,\phi))^2$ are defined with full covariance structures. However, for computational simplicity, the numerical example in Section 5 assumes uncorrelated inputs and outputs. In this section, we investigate how the estimated input vector changes when different covariance structures are introduced.
We consider three covariance scenarios for the inputs and outputs of DMU01:

- Scenario A (Base case): All inputs and outputs are uncorrelated (as in the original analysis).
- Scenario B (Positive correlation): Inputs are positively correlated with correlation coefficient $\rho = 0.3$ among all pairs of inputs; outputs are similarly positively correlated with $\rho = 0.3$ among all pairs of outputs.
- Scenario C (Mixed correlation): Personnel rate and payable profits are positively correlated ($\rho = 0.5$), while delayed requisitions are negatively correlated with both other inputs ($\rho = -0.2$). Outputs follow a similar mixed pattern.

The correlation structure is incorporated into the covariance matrices in the variance terms $(\sigma_i^I(\mu,\theta))^2$ and $(\sigma_r^O(\mu,\phi))^2$ of models (1)-(17), which are then transformed into the deterministic equivalents (17)-(18) and subsequently into the SMOP model (21).

Table 12 presents the estimated input vectors for DMU01 under a 10% efficiency improvement target at $\alpha = 0.01$ for the three covariance scenarios. All values reported in Table 12 are expressed as percentage variations relative to the original input values, $\mathbf{X}_{01}$.

Some important observations from the results in Table 12 :

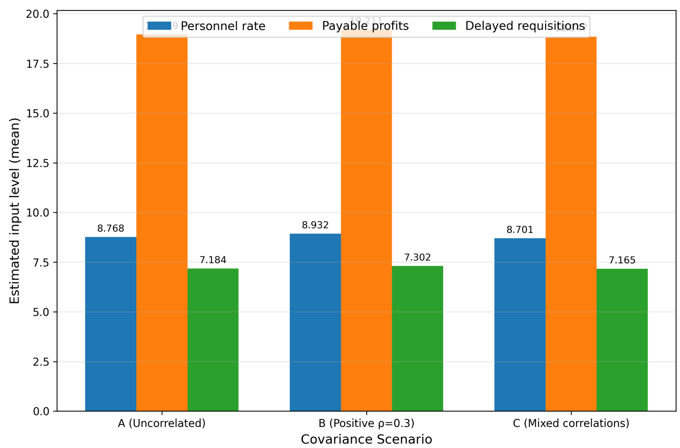

Figure 5: Input variation across covariance scenarios

1. Covariance matters: The estimated input vector changes noticeably when covariance terms are included. Under positive correlation (Scenario B), the required reductions are less aggressive (or even require increases for some inputs) compared to the uncorrelated case. This occurs because positive correlation amplifies the effective uncertainty, making it more difficult to satisfy the probabilistic constraints with the same level of confidence.

2. Mixed correlations reduce uncertainty: Scenario C (mixed correlations) yields slightly more aggressive reductions than the base case. Negative correlation between delayed requisitions and other inputs reduces the overall uncertainty in the convex combinations, allowing the model to propose larger reductions while still maintaining the 99% confidence requirement.

3. Managerial implications: The sensitivity to covariance structure underscores the importance of accurately estimating correlations when implementing the proposed SMOP model. Banks with positively correlated inputs and outputs should be more conservative in their resource reduction targets, while those with mixed correlations (where risks offset each other) can be more aggressive. This provides a quantitative basis for risk-adjusted resource allocation.

4. Robustness of the approach: Despite the variations in the estimated values, the relative ranking of the three inputs remains consistent across all scenarios: personnel rate shows the largest reduction (or smallest increase), delayed requisitions shows the smallest reduction (or largest increase), and payable profits fall in between. This suggests that the qualitative conclusions from our analysis are robust to reasonable variations in the covariance structure.

5. Validation of the SMOP formulation: The fact that the model continues to produce feasible and interpretable solutions under all three covariance scenarios validates the SMOP formulation's ability to handle general covariance structures.

Profound risk-structure interactions are invisible when examining means alone. Extending the covariance analysis to the variance dimension reveals this risk. In Scenario B, not only do the means increase (making reductions harder to achieve), but the variances of all three inputs increase dramatically by 132% for personnel, 150.85% for payable profits, and 41.38% for delayed requisitions. This amplification occurs because positive correlations reinforce the joint uncertainty of the convex combinations, making the stochastic constraints much harder to satisfy at the same confidence level. Consequently, the model must "loosen" all inputs (increasing both their expected levels and their volatility) to maintain feasibility at $\alpha = 0.01$.

Conversely, Scenario C generates the most desirable risk profile. The variance of delayed requisitions actually decreases by -3.45% relative to the original data, while the mean is reduced by -0.87%. This unique outcome –simultaneously reducing expected levels and volatility– occurs because negative correlations between delayed requisitions and the other inputs create a natural hedging effect. For bank managers, this insight is highly actionable: actively managing the portfolio of inputs to introduce negative correlations (e.g.,, incentivizing loan officers to balance high-risk, high-return loans with low-risk, low-return deposits) can reduce both expected resource consumption and its associated uncertainty. Thus, covariance structure analysis, explicitly incorporating variance changes, enables a dual optimization of efficiency improvement and risk minimization, providing a truly comprehensive decision-support tool that parallels the mean-variance insights established in Tables 5-7.

6 Conclusions, limitations, and future research directions

In this study, a stochastic decision-making environment was assumed in which a central unit (the decision-maker) controls the resources and products of a set of DMUs. SDEA and SMOP problems were incorporated to allow a manager to combine priority information in the stochastic framework. The InvDEA problems were extended for resources/products estimation under stochastic data. Stochastic (weak) Pareto solutions were utilized in the significance level α of SMOP problems to extract the necessary and sufficient conditions for resources/products estimation. It is worth noting that through the several objective models, the decision-maker can find the most-preferred resource allocation/investment analysis solution by verifying tradeoffs between the whole efficiency at the global organizational level and equity between all DMUs.The proposed models accommodate arbitrary variations in resource/product levels-including increases, decreases, and unchanged values-while simultaneously improving the efficiency score at a specified significance level. The application of the proposed InvDEA technique in the banking sector was discussed to achieve a given efficiency target. The numerical results demonstrated how a branch manager can:

- Determine which resources to increase or decrease to achieve a specific efficiency improvement (e.g.,, 10%, 20%, or full efficiency);
- Interpret efficiency scores across different significance levels (Table 3);
- Use sensitivity analysis on α (Section 5.3, Figure 3) to balance risk and resource reallocation aggressiveness;
- Explore trade-offs among inputs via the Pareto frontier.

Also, the numerical application in this study uses a dataset of 20 bank branches. While this sample size is comparable to many empirical DEA studies in the banking sector (see, e.g.,, Behzadi & Mirbolouki [5] and Ghomi et al. [28]), it is important to acknowledge that small samples can affect the reliability of efficiency estimates in SDEA for several reasons. First, the estimation of variances and covariances becomes less precise with fewer observations, which may propagate into the chance constraints and lead to biased efficiency scores. Second, the discriminatory power of DEA decreases as the number of DMUs approaches the number of inputs and outputs, although our dataset (20 DMUs, 3 inputs, 5 outputs) comes close to the common rule of thumb that the number of DMUs should be at least three times the total number of inputs and outputs (i.e.,, $3 \times (3+5) = 24$, whereas we have 20). Third, the asymptotic properties of chance-constrained SDEA models rely on large-sample approximations; with small samples, the estimated efficiency scores may exhibit higher variability. To mitigate these concerns, we used variance and covariance estimates adopted from Behzadi and Mirbolouki [5], which were derived from the same dataset, ensuring consistency. For practical implementation, decision-makers are encouraged to use larger datasets whenever available, or to employ bootstrap or jackknife techniques to assess the stability of the estimated efficiency scores. Future research could investigate the finite-sample properties of the proposed SMOP-based InvDEA models through simulation studies across various sample sizes.

The proposed stochastic InvDEA framework has several limitations that should be acknowledged. First, the models rely on the assumption that all random inputs and outputs follow a multivariate normal distribution (or, equivalently, a symmetric error structure). Although this assumption is widely adopted in stochastic DEA and enables the derivation of tractable deterministic equivalents, it may not adequately represent real-world data that exhibit skewness, heavy tails, or other departures from normality. Extending the proposed framework to accommodate alternative probability distributions, such as log-normal, gamma, or other asymmetric distributions, as well as copula-based models for capturing dependence among non-normal variables, would considerably broaden its applicability.

Second, the proposed models assume that DMUs are independent of one another. In many practical applications, however, operational or environmental interactions may exist among units, and ignoring such dependencies could affect the accuracy of efficiency evaluation and target setting. Developing stochastic inverse DEA models that explicitly account for interdependencies among DMUs therefore represents an important avenue for future research.

Third, the proposed models are developed under the assumption of CRS. The CRS assumption is adopted to measure overall technical efficiency and to focus on the proportional relationship between resource utilization and output generation. In the context of banking, this assumption can be considered reasonable in the long run, where banking operations may be scaled proportionally with changes in resource utilization. However, the CRS assumption may not fully capture production technologies characterized by VRS . Therefore, extending the proposed framework to accommodate VRS or other production technologies would further enhance its flexibility and practical relevance.

Fourth, The proposed SMOP framework requires estimates of the covariance matrices among inputs and among outputs. In practice, these matrices must be estimated from historical data. When sample sizes are limited (as in our application with 20 branches), several practical approaches can be employed: Pooled covariance estimation, Shrinkage estimators, Factor models, Bootstrap resampling, and Scenario-based sensitivity analysis.

It is worth noting that, in our application, we used variance and covariance estimates adopted from Behzadi and Mirbolouki [5], which were derived from the same dataset. For new applications, we recommend that practitioners combine the above approaches based on data availability, and complement them with the sensitivity analysis presented in Section 5.5 to assess the robustness of their conclusions.

Finally, as the number of DMUs, inputs, outputs, or stochastic constraints increases, the computational burden of solving the stochastic multi-objective programming models also increases. Conse-

Table 13: The inputs of 20 bank branches

	Input1		Input2		Input3	
DMU	Mean	Variance	Mean	Variance	Mean	Variance
DMU01	9.131	0.05	18.79	8.81	7.228	0.58
DMU02	10.59	0.53	44.32	24.1	1.121	0.02
DMU03	6.712	0.86	19.73	27.7	19.21	0.47
DMU04	11.91	0.31	17.43	12.2	59.47	5.99
DMU05	7.012	0.02	10.38	2.12	12.23	0.85
DMU06	18.99	0.88	16.67	10.8	568.6	28.1
DMU07	11.16	0.01	25.46	18.6	552.8	43.2
DMU08	15.05	0.48	123.1	42.6	14.78	0.06
DMU09	8.787	0.38	36.16	38.4	361.8	23.2
DMU10	19.88	0.25	46.41	53.1	12.81	0.38
DMU11	18.92	0.17	36.88	54.5	24.43	0.01
DMU12	20.45	0.42	100.8	31.8	115.2	19.4
DMU13	12.41	0.12	20.19	10.6	78.02	24.1
DMU14	8.051	0.79	33.21	24.3	115.3	15.6
DMU15	18.48	0.92	45.36	92.6	57.52	12.8
DMU16	10.35	0.27	11.16	3.32	43.32	36.1
DMU17	9.511	0.01	31.49	38.5	173.3	3.13
DMU18	13.71	0.18	40.32	51.4	10.88	0.12
DMU19	11.69	0.26	26.44	26.2	31.22	0.05
DMU20	7.823	0.58	17.74	10.1	13.06	8.88

quently, developing efficient decomposition techniques, approximation algorithms, or metaheuristic solution methods for large-scale problems constitutes another promising direction for future research.

Beyond these methodological extensions, the proposed framework may also be generalized to more sophisticated DEA structures, including dynamic DEA and network DEA models, thereby enabling efficiency improvement and target setting in multi-period and multi-stage production systems under uncertainty.

Appendix A

Source: The data used in this study were adopted from Behzadi and Mirbolouki [5] and are presented in Tables 13 and 14. The corresponding variance values were also adopted from the same source.

Acknowledgements

The authors sincerely thank the Editor and the anonymous reviewers of JMM for their careful evaluation and constructive comments, which have helped improve the quality and clarity of the manuscript.

Declaration

The authors declare no competing interests.

Table 14: The outputs of 20 bank branches

	output1		output2		output3		output4		output5	
DMU	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance	Mean	Variance
DMU01	149.85	48.51	49.621	18.01	4.701	0.41	4.748	0.41	30.09	24.34
DMU02	50.772	8.011	73.132	13.11	1.815	0.02	3.035	0.68	5.823	0.689
DMU03	259.91	29.59	108.04	22.56	6.016	0.01	10.06	2.98	2.721	0.4392
DMU04	137.51	21.65	44.972	13.78	4.923	1.65	4.212	3.84	63.61	33.73
DMU05	95.901	2.521	31.633	18.94	2.718	0.44	9.024	7.23	64.55	31.25
DMU06	112.58	3.562	71.958	17.05	13.19	9.15	41.89	13.2	291.3	92.52
DMU07	192.97	14.56	78.015	19.56	7.791	3.56	15.89	2.34	7.499	1.268
DMU08	724.38	66.02	219.69	37.56	35.32	15.3	23.98	10.4	361.6	48.35
DMU09	548.15	41.86	86.225	28.31	17.64	3.67	86.23	17.2	565.2	175.1
DMU10	1229.1	69.02	194.58	17.35	25.91	12.3	86.76	22.1	600.6	86.34
DMU11	11557	718.1	155.32	49.13	166.6	14.1	8.142	3.11	119.9	14.89
DMU12	1132.1	35.35	248.16	23.89	46.88	26.3	31.85	4..1	96.21	44.47
DMU13	438.39	17.41	104.41	25.71	10.68	1.04	30.22	3.77	331.9	17.16
DMU14	260.82	15.32	87.369	23.44	8.415	3.52	6.101	2.86	36.93	5.777
DMU15	11190	1214	166.44	10.69	65.12	16.1	132.7	13.9	919.1	133.1
DMU16	709.85	42.19	159.48	32.78	36.89	6.28	12.15	2.51	79.66	10.52
DMU17	308.11	43.72	107.03	29.39	11.87	1.76	13.63	1.57	342.3	134.9
DMU18	259.21	34.21	81.779	21.65	5.212	1.94	8.021	3.52	107.9	13.07
DMU19	381.33	57.38	72.993	29.26	5.165	1.79	50.32	7.47	577.2	113.4
DMU20	399.25	10.41	40.985	15.93	11.51	0.99	6.432	3.25	82.13	9.891

References

[1] O. Abdolazimi, M. Sepehrifar, D. Shishebori, I. Banicescu, S. Azad, K. Pishini, J. Ma, *A stochastic DEA-GARMA integration for dynamic forecasting of undesirable outputs applied to the temporal efficiency analysis of water and sewerage services*, J. Clean. Prod. **521** (2025) 146177.

[2] G.R. Amin, S. Al-Muharrami, *A new inverse data envelopment analysis model for mergers with negative data*, IMA J. Manag. Math. **29** (2018) 137–149.

[3] G.R. Amin, M. Ibn Boamah, *A two-stage inverse data envelopment analysis approach for estimating potential merger gains in the US banking sector*, Manag. Decis. Econ. **42** (2021) 2112–2124.

[4] B. Arabi, M. Toloo, H. Farnoudkia, B. Xu, *Introducing stochastic data envelopment analysis models for evaluating circular economy performance in European countries: pathways towards sustainability*, OR Spectr. **48** (2026) 185–222.

[5] M.H. Behzadi, M. Mirbolouki, *Symmetric error structure in stochastic DEA*, Int. J. Ind. Math. **4** (2012) 335–343.

[6] M.H. Behzadi, N. Nematollahi, M. Mirbolouki, *Ranking efficient DMUs with stochastic data by considering inefficient frontier*, Int. J. Ind. Math. **1** (2009) 219–226.

[7] V.J. Bolós, R. Benítez, V. Coll-Serrano, *Chance constrained directional models in stochastic data envelopment analysis*, Oper. Res. Perspect. **12** (2024), 100307.

[8] V.J. Bolós, V. Coll, R. Benítez, *SdeaR: Stochastic data envelopment analysis with R*, SSRN Working Paper (2026).

[9] A. Charnes, W.W. Cooper, E. Rhodes, *Measuring the efficiency of decision making units*, Eur. J. Oper. Res. **2** (1978) 429–444.

[10] W.W. Cooper, H. Deng, Z.M. Huang, S X. Li, *Change constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis*, J. Oper. Res. Soc. **53** (2002) 1347–1356.

[11] W.W. Cooper, Z. Huang, V. Lelas, S. Li, O.B. Olesen, *Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA*, J. Prod. Anal. **9** (1998) 53–80.

[12] W.W. Cooper, Z. Huang, S. Li, *Satisficing DEA models under chance constraints*, Ann. Oper. Res. **66** (1996) 279–295.

[13] L. Deng, C. Li, *A fuzzy stochastic DEA model considering an input–output structure*, Axioms **15** (2026) 376.

[14] D.J. Lee, *Inverse DEA with frontier changes for new target setting*, Eur. J. Oper. Res. **254** (2016) 510–516.

[15] A. Emrouznejad, G. R. Amin, *Advances in inverse data envelopment analysis: empowering performance assessment*, IMA J. Manag. Math. **34** (2023) 415–419.

[16] A. Emrouznejad, G.R. Amin, M. Ghiyasi, M. Michali, *A review of inverse data envelopment analysis: origins, development and future directions*, IMA J. Manag. Math. **34** (2023) 421–440.

[17] A. Emrouznejad, M. Tavana, *Performance Measurement with Fuzzy Data Envelopment Analysis*, Springer, Berlin, 2014.

[18] A. Emrouznejad, G.-L. Yang, G.R. Amin, *A novel inverse DEA model with application to allocate the CO_2 emissions quota to different regions in Chinese manufacturing industries*, J. Oper. Res. Soc. **70** (2019) 1079–1090.

[19] S. Gattoufi, G.R. Amin, A. Emrouznejad, *A new inverse DEA method for merging banks*, IMA J. Manag. Math. **25** (2014) 73–87.

[20] M. Ghiyasi, *Inverse DEA based on cost and revenue efficiency*, Comput. Ind. Eng. **114** (2017) 258–263.

[21] M. Ghiyasi, N. Nasrabadi, A. Emrouznejad, *Inverse DEA based on cost and revenue efficiency in the absence of price information*, Expert Syst. Appl. **284** (2025) 128028.

[22] M. Ghiyasi, A. Emrouznejad, G.R. Amin, *Generalizing inverse data envelopment analysis through directional distance function*, Expert Syst. Appl. **299** (2026) 129996.

[23] S. Ghobadi, *A generalized DEA model for inputs (outputs) estimation under inter-temporal dependence*, RAIRO Oper. Res. **53** (2019) 1791–1805.

[24] S. Ghobadi, *A dynamic DEA model for resource allocation*, Int. J. Math. Oper. Res. **17** (2020) 50–77.

[25] S. Ghobadi, K. Soleimani-Chamkhorami, *Merging decision-making units with fuzzy data*, Asia-Pacific J. Oper. Res. **39** (2022) 2140012.

[26] S. Ghobadi, K. Soleimani-Chamkhorami, E. Zanboori, *A novel inverse DEA model for restructuring DMUs with negative data*, Int. J. Oper. Res. **46** (2023) 118–132.

[27] S. Ghobadi, E. Zanboori, K. Soleimani-Chamkhorami, M. Mahmoudi, *Restructuring evaluation using generalized inverse DEA-R frameworks*, TOP **34** (2026) 136–154.

[28] A. Ghomi, S. Ghobadi, M.H. Behzadi, M. Rostamy-Malkhalifeh, *Inverse data envelopment analysis with stochastic data*, RAIRO Oper. Res. **55** (2021) 2739–2762.

[29] A. Hadi-Vencheh, A.A. Foroughi, M. Soleimani-Damaneh, *A DEA model for resource allocation*, Econom. Model. **25** (2008) 983–993.

[30] F. Hosseinzadeh Lotfi, N. Nematollahi, M. H. Behzadi, M. Mirbolouki, Z. Moghaddas, *Centralized resource allocation with stochastic data*, J. Comput. Appl. Math. **236** (2012) 1783–1788.

[31] F. Hosseinzadeh Lotfi, N. Nematollahi, M. H. Behzadi, M. Mirbolouki, *Ranking decision making units with stochastic data by using coefficient of variation*, Math. Comput. Appl. **15** (2010) 148–155.

[32] Z. Huang, S.X. Li, *Stochastic DEA models with different types of input-output disturbances*, J. Prod. Anal. **15** (2001) 95–113.

[33] G.R. Jahanshahloo, F.H. Lotfi, N. Shoja, G. Tohidi, S. Razavyan, *Input estimation and identification of extra inputs in inverse DEA models*, Appl. Math. Comput. **156** (2004) 427–437.

[34] G.R. Jahanshahloo, F. H. Lotfi, N. Shoja, G. Tohidi, S. Razavyan, *The outputs estimation of a DMU according to improvement of its efficiency*, Appl. Math. Comput. **147** (2004) 409–413.

[35] G.R. Jahanshahloo, M.H. Behzadi, M. Mirbolouki, *Ranking stochastic efficient DMUs based on reliability*, Int. J. Ind. Math. **2** (2010) 263–270.

[36] G.R. Jahanshahloo, M. Soleimani-Damaneh, S. Ghobadi, *Inverse DEA under inter-temporal dependence using multiple-objective programming*, Eur. J. Oper. Res. **240** (2015) 447–456.

[37] M. Khodabakhshi, *Estimating most productive scale size with stochastic data in data envelopment analysis*, Econom. Model. **26** (2009) 968–973.

[38] S.X. Li, *Stochastic models and variable returns to scales in data envelopment analysis*, Eur. J. Oper. Res. **104** (1998) 532–548.

[39] S.-W. Lin, W.-M. Lu, *A comparison of chance-constrained data envelopment analysis, stochastic nonparametric envelopment of data and bootstrap method: A case study of cultural regeneration performance of cities*, Eur. J. Oper. Res. **316** (2024) 1179–1191.

[40] M. Majid Kalantary, R. F. Saen, *Assessing sustainability of supply chains: An inverse network dynamic DEA model*, Comput. Ind. Eng. **126** (2018) 272–287.

[41] P. Peykani, E. Mohammadi, M. S. Pishvaee, M. Rostamy-Malkhalifeh, A. Jabbarzadeh, *A novel fuzzy data envelopment analysis based on robust possibilistic programming: Possibility, necessity and credibility-based approaches*, RAIRO Oper. Res. **52** (2018) 1445–1463.

[42] P. Peykani, E. Mohammadi, R.F. Saen, S.J. Sadjadi, M. Rostamy-Malkhalifeh, *Data envelopment analysis and robust optimization: A review*, Expert Syst. **37** (2020) e12534.

[43] J.K. Sengupta, *Efficiency analysis by stochastic data envelopment analysis*, Appl. Econ. Lett. **7** (2000) 379–383.

[44] L. Simar, *How to improve the performances of DEA/FDH estimators in the presence of noise*, J. Prod. Anal. **28** (2007) 183–201.

[45] L. Simar, P.W. Wilson, *A general methodology for bootstrapping in non-parametric frontier models*, J. Appl. Stat. **27** (2000) 779–802.

[46] K. Soleimani-Chamkhorami, F. Hosseinzadeh Lotfi, G.R. Jahanshahloo, M. Rostamy-Malkhalifeh, *A ranking system based on inverse data envelopment analysis*, IMA J. Manag. Math. **31** (2020) 367–385.

[47] Kh. Soleimani-Chamkhorami, F. Hosseinzadeh Lotfi, G.R. Jahanshahloo, M. Rostamy-Malkhalifeh, *Preserving cost and revenue efficiency through inverse data envelopment analysis models*, INFOR: Inf. Syst. Oper. Res. **58** (2020) 561–578.

[48] S. Thore, *Chance-constrained activity analysis*, Eur. J. Oper. Res. **30** (1987) 267–269.

[49] M. Wegener, G.R. Amin, *Minimizing greenhouse gas emissions using inverse DEA with an application in oil and gas*, Expert Syst. Appl. **122** (2019) 369–375.

[50] Q.L. Wei, J. Z. Zhang, X.S. Zhang, *An inverse DEA model for inputs/outputs estimate*, Eur. J. Oper. Res. **121** (2000) 151–163.

[51] E. Zeinodin, S. Ghobadi, *Merging DMUs based on of the idea inverse DEA*, Iran. J. Optim. **12** (2019) 77–84.

[52] G. Zhang, J. Cui, *A general inverse DEA model for non-radial DEA*, Comput. Ind. Eng. **142** (2020) 106368.

[53] J. Zhu, *Imprecise DEA via standard linear DEA models with a revisit to a Korean mobile telecommunication company*, Oper. Res. **52** (2004) 323–329.

[54] K. Miettinen, *Nonlinear Multiobjective Optimization*, Kluwer Academic Publishers, Boston, 1999.