Bolós, Vicente J.; Benítez, Rafael; Coll-Serrano, V.

Article

Chance constrained directional models in stochastic data envelopment analysis

Operations Research Perspectives

Provided in Cooperation with:
Elsevier

Suggested Citation: Bolós, Vicente J.; Benítez, Rafael; Coll-Serrano, V. (2024) : Chance constrained directional models in stochastic data envelopment analysis, Operations Research Perspectives, ISSN 2214-7160, Elsevier, Amsterdam, Vol. 12, pp. 1-8, https://doi.org/10.1016/j.orp.2024.100307

This Version is available at:
https://hdl.handle.net/10419/325789

Operations Research Perspectives 12 (2024) 100307

Contents lists available at ScienceDirect

Operations Research Perspectives

journal homepage: www.elsevier.com/locate/orp

Chance constrained directional models in stochastic data envelopment analysis

V.J. Bolós [a,*], R. Benítez [a], V. Coll-Serrano [b]

[a] Dpto. Matemáticas para la Economía y la Empresa, Facultad de Economía., Universidad de Valencia. Avda. Tarongers s/n, 46022 Valencia, Spain
[b] Dpto. Economía Aplicada, Facultad de Economía., Universidad de Valencia. Avda. Tarongers s/n, 46022 Valencia, Spain

ARTICLE INFO

Keywords:
Data envelopment analysis
Stochastic DEA
Chance constrained DEA
Efficiency
Directional models

ABSTRACT

We construct a new family of chance constrained directional models in stochastic data envelopment analysis, generalizing the deterministic directional models and the chance constrained radial models. We prove that chance constrained directional models define the same concept of stochastic efficiency as the one given by chance constrained radial models and, as a particular case, we obtain a stochastic version of the generalized Farrell measure. Finally, we give some examples of application of chance constrained directional models with stochastic and deterministic directions, showing that inefficiency scores obtained with stochastic directions are less or equal than those obtained considering deterministic directions whose values are the means of the stochastic ones.

1. Introduction

Data envelopment analysis (DEA) [1] is a non-parametric technique used to measure the relative efficiency of a homogeneous set of decision-making units (DMUs) that use multiple inputs to obtain multiple outputs. Using mathematical programming methods, DEA allows the identification of the best practice frontier of the production possibility set, determined by DMUs which are qualified as efficient. On the other hand, DMUs that move away from this frontier are considered inefficient.

Radial models were introduced in [1–3] for constant returns to scale (CCR models), and [4] for variable returns to scale (BCC models). In general, they can be either input or output-oriented. In the former case, we look for determining the maximal proportionate reduction of inputs allowed by the production possibility set, while maintaining the current output level. On the other hand, in the output-oriented case, we want to find the maximal proportionate increase of outputs while keeping the current input consumption.

One of the drawbacks of radial models is that these reductions (of inputs) or increases (of outputs) have to be made in the same proportion. Therefore, [5] defined non-radial models that allow non-proportional movements in inputs and outputs. Subsequently, [6] introduced the so-called additive models, which are also non-radial and do not distinguish between orientations. Later, [7,8] constructed the slacks-based measure of efficiency (SBM) models, with the same characteristics as additive models but providing an efficiency score. However, all these models compute efficiency with respect to the target located at the "farthest" point within the portion of the efficient frontier that dominates the evaluated DMU and, therefore, may be inappropriate in the case of wanting to calculate a "close" efficient target. Some authors adapted these non-radial models to the search of the closest targets. For example, [9] adapted several models and constructed the mADD model as the minimum-distance version of the additive model. In turn, [10] developed the SBM-max model. Nevertheless, these models are not weakly monotonic in general, in the sense that a DMU can get a worse efficiency score than another DMU it outperforms.

Other models allowing non-proportional movements in inputs and outputs were introduced in [11]. Later, the directional models were developed in [12–14]. These models, apart from being weakly monotonic, generalize the classical radial models and can measure efficiency in the full input–output space. The main goal of directional models is that the user can change the proportion in which the inputs and outputs are modified, thus defining a custom orientation taking account of particularities of the market and characterizing the criteria of management chosen by the producer.

Nevertheless, all these models assume a deterministic behavior of the inputs and outputs of the DMUs, and are not appropriate in a stochastic framework. In this sense, stochastic DEA was developed in two main directions (see [15] for a complete review on stochastic DEA). On one hand, [16] initiated one approach that included statistical axioms defining a statistical model and a sampling process into the DEA framework. On the other hand, chance constrained DEA [17,18]

* Corresponding author.
E-mail addresses: vicente.bolos@uv.es (V.J. Bolós), rabesua@uv.es (R. Benítez), vicente.coll@uv.es (V. Coll-Serrano).

https://doi.org/10.1016/j.orp.2024.100307
Received 4 March 2024; Received in revised form 13 May 2024; Accepted 7 June 2024
Available online 10 June 2024

allows for uncertainty in the inputs and outputs, ensuring that the efficiency scores are obtained with a given degree of confidence and hence, providing a more robust and realistic assessment of efficiency compared to deterministic DEA. Chance constrained versions of some radial models were developed in [19–21]. More recently, in [22], a directional chance constrained model was applied to the very particular case of portfolios, in which there is one input and one output that are dependent on the portfolio return and its probability distribution. However, up to our knowledge, there are no more chance constrained directional models in the literature. Therefore, in the same way that directional models generalize the classical radial models in the deterministic DEA, the objective of this work is to define chance constrained directional models, generalizing both chance constrained radial models, and deterministic directional models. In this way, as we will see below, the proposed models will allow us to select the proportion in which inputs and outputs are modified to achieve efficiency within a stochastic framework.

Finally, it is worth mentioning that fuzzy DEA is another recent approach to DEA analysis presenting uncertainty in inputs and outputs, with Kao–Liu [23], Guo–Tanaka [24] and possibilistic [25] models as the most relevant. More recently, some chance constrained fuzzy DEA models has been developed in [26]. However, although fuzzy DEA methods have proven to be very useful for the analysis of data where uncertainty arises from ambiguity in the variables (as is the case for linguistic variables), these methods do not make use of the information derived from the probability distribution of the data.

The paper is organized as follows. In Section 2, we review directional models and adapt the notation to be able to define, in Section 3, their chance constrained versions in which inputs and outputs become stochastic. We have developed two types of chance constrained directional models: with *stochastic directions* and with *deterministic directions*. Models with stochastic directions are introduced in Section 3.1 and they generalize the existing chance constrained radial models. Moreover, they serve to define a stochastic version of the generalized Farrell measure. On the other hand, models with deterministic directions are defined analogously in Section 3.2. Next, in Section 3.3, we explore the joint chance constrained directional models assuming that variables are stochastically independent. In Section 4, we give some examples of application of chance constrained directional models with stochastic or deterministic directions, and, finally, we present some conclusions in Section 5, raising some open issues.

In general, we denote vectors by bold-face letters and they are considered as column vectors unless otherwise stated. The elements of a vector are denoted by the same letter as the vector, but unbolded and with subscripts. The 0-vector is denoted by $\mathbf{0}$ and the context determines its dimension.

2. Directional models

We consider $D = \{\mathrm{DMU}_1, \dots, \mathrm{DMU}_n\}$ a set of n DMUs with m inputs and s outputs. Matrices $X = (x_{ij})$ and $Y = (y_{rj})$ are the *input* and *output data matrices*, respectively, where x_{ij} and y_{rj} denote the ith input and rth output of the jth DMU. We also assume that x_{ij} and y_{rj} are all strictly positive. Given $\mathrm{DMU}_o \in D$, we define the column vectors $\mathbf{x}_o = (x_{1o}, \dots, x_{mo})^\top$ and $\mathbf{y}_o = (y_{1o}, \dots, y_{so})^\top$. Although we are going to consider constant returns to scale (CRS) in all programs, different returns to scale can be easily considered by adding the corresponding constraints: $\mathbf{e}\lambda = 1$ for variable returns to scale (VRS), $0 \le \mathbf{e}\lambda \le 1$ for non-increasing returns to scale (NIRS), $\mathbf{e}\lambda \ge 1$ for non-decreasing returns to scale (NDRS) or $L \le \mathbf{e}\lambda \le U$ for generalized returns to scale (GRS), with $0 \le L \le 1$ and $U \ge 1$, where $\lambda \in \mathbb{R}^n$ and $\mathbf{e} = (1, \dots, 1) \in \mathbb{R}^n$ is the all-ones row vector.

Directional models were introduced in [13,14]. Given $\mathrm{DMU}_o \in D$, the associated linear program under CRS and its second stage are given by

$$\text{(a)}\ \begin{aligned}\max_{\beta,\lambda}\quad & \beta\\ \text{s.t.}\quad & \beta\mathbf{g}^- + X\lambda \le \mathbf{x}_o,\\ & -\beta\mathbf{g}^+ + Y\lambda \ge \mathbf{y}_o,\\ & \lambda \ge \mathbf{0},\end{aligned} \qquad \text{(b)}\ \begin{aligned}\max_{\lambda,\mathbf{s}^-,\mathbf{s}^+}\quad & \mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\\ \text{s.t.}\quad & X\lambda + \mathbf{s}^- = \mathbf{x}_o - \beta^*\mathbf{g}^-,\\ & Y\lambda - \mathbf{s}^+ = \mathbf{y}_o + \beta^*\mathbf{g}^+,\\ & \lambda \ge \mathbf{0},\ \mathbf{s}^- \ge \mathbf{0},\ \mathbf{s}^+ \ge \mathbf{0},\end{aligned} \tag{1}$$

where $\lambda \in \mathbb{R}^n$ and the *slacks* $\mathbf{s}^- \in \mathbb{R}^m$ and $\mathbf{s}^+ \in \mathbb{R}^s$ are non-negative column vectors, $\mathbf{g} = (-\mathbf{g}^-, \mathbf{g}^+) \ne \mathbf{0}$ is a preassigned *direction* (with $\mathbf{g}^- \in \mathbb{R}^m$ and $\mathbf{g}^+ \in \mathbb{R}^s$ non-negative column vectors), while the *weights* $\mathbf{w}^- \in \mathbb{R}^m$ and $\mathbf{w}^+ \in \mathbb{R}^s$ are strictly positive row vectors. Moreover, in (1) (b), β^* is the optimal objective value of the first stage program (1) (a), which is always greater than or equal to 0.

The advantage of directional models is that the way DMU_o is projected onto the efficient frontier can be customized by using a direction $\mathbf{g}$ according to the criteria of management chosen by the producer [12].

Definition 2.1. $\mathrm{DMU}_o \in D$ is *efficient* if and only if the optimal objectives β^* and $\mathbf{w}^-\mathbf{s}^{-*} + \mathbf{w}^+\mathbf{s}^{+*}$ in (1) are 0, where $\mathbf{s}^{-*}, \mathbf{s}^{+*}$ are optimal slacks. Moreover, we say that DMU_o is *weakly efficient* if $\beta^* = 0$ and it is not efficient.

Remark 2.1. The concept of efficiency given in Definition 2.1 does not depend on the direction $\mathbf{g}$ because $\beta^* = 0$ and hence, according to (1), the optimal solution remains optimal if we change the direction. So, if DMU_o is efficient for a given direction, then it is efficient for any direction.

Remark 2.2. We consider that the weights $\mathbf{w}^-$ and $\mathbf{w}^+$ are strictly positive and hence, $\mathbf{w}^-\mathbf{s}^{-*} + \mathbf{w}^+\mathbf{s}^{+*} = 0$ if and only if $\mathbf{s}^{-*} = \mathbf{0}$ and $\mathbf{s}^{+*} = \mathbf{0}$. However, allowing zero weights is useful in some cases. For example, we can introduce *non-discretionary* inputs and outputs (which are exogenously fixed, see [27, Chapter 7]) by setting the corresponding directions and weights $g_i^- = w_i^- = 0$ (for the ith input) or $g_r^+ = w_r^+ = 0$ (for the rth output). In this case, the optimal slacks of non-discretionary variables are not taken into account in the efficiency evaluation.

Nevertheless, in order to construct the chance constrained version of the model in Section 3, it is convenient to write the second stage program as

$$\begin{aligned}\max_{\lambda,\mathbf{s}^-,\mathbf{s}^+}\quad & \mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\\ \text{s.t.}\quad & X\lambda + \mathbf{s}^- \le \mathbf{x}_o - \beta^*\mathbf{g}^-,\\ & Y\lambda - \mathbf{s}^+ \ge \mathbf{y}_o + \beta^*\mathbf{g}^+,\\ & \lambda \ge \mathbf{0},\ \mathbf{s}^- \ge \mathbf{0},\ \mathbf{s}^+ \ge \mathbf{0}.\end{aligned} \tag{2}$$

Obviously, programs (1) (b) and (2) are equivalent, since any optimal solution of (2) should also satisfy the constraints of (1) (b), and vice versa.

The one-stage version of (1) (a) and (2) is given by

$$\begin{aligned}\max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+}\quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right)\\ \text{s.t.}\quad & \beta\mathbf{g}^- + X\lambda + \mathbf{s}^- \le \mathbf{x}_o,\\ & -\beta\mathbf{g}^+ + Y\lambda - \mathbf{s}^+ \ge \mathbf{y}_o,\\ & \lambda \ge \mathbf{0},\ \mathbf{s}^- \ge \mathbf{0},\ \mathbf{s}^+ \ge \mathbf{0},\end{aligned} \tag{3}$$

where ε a positive non-Archimedean infinitesimal. Although the use of non-Archimedean infinitesimals is not numerically rigorous, in the case of DEA models with two stages is very useful because it synthesizes the two stages in a single program and you can always do the reverse process to recover the programs of each stage.

Let us consider the directions $\mathbf{g}^- = D^-\mathbf{x}_o$ and $\mathbf{g}^+ = D^+\mathbf{y}_o$, where the matrices $D^- = \mathrm{diag}(d_1^-, \dots, d_m^-)$ and $D^+ = \mathrm{diag}(d_1^+, \dots, d_s^+)$ are diagonal

with $d_1^-, \dots, d_m^-, d_1^+, \dots, d_s^+ \geq 0$. Then (3) becomes

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon \left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & \Theta^-(\beta)\mathbf{x}_o - X\lambda - \mathbf{s}^- \geq \mathbf{0}, \\ & \Theta^+(\beta)\mathbf{y}_o - Y\lambda + \mathbf{s}^+ \leq \mathbf{0}, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{4}$$

where

$$\Theta^-(\beta) = I_m - \beta D^-, \qquad \Theta^+(\beta) = I_s + \beta D^+, \tag{5}$$

with I_m, I_s the $m \times m$ and $s \times s$ identity matrices, respectively. The optimal value β^* given by (4) coincides with the *oriented Farrell proportional distance* defined by [12], with *orientation* $(\mathbf{d}^-, \mathbf{d}^+)$, where $\mathbf{d}^- = (d_1^-, \dots, d_m^-)$ and $\mathbf{d}^+ = (d_1^+, \dots, d_s^+)$. Hence, by Definition 2.1, DMU_o is efficient if and only if the optimal objective value of (4) is 0, regardless of the orientation.

As a particular case, if $D^- = I_m$ (i.e. $d_1^-, \dots, d_m^- = 1$), and $D^+ = 0$ (i.e. $d_1^+, \dots, d_s^+ = 0$), then we can write (4) as

$$\begin{aligned} \min_{\theta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \theta - \varepsilon \left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & \theta\mathbf{x}_o - X\lambda - \mathbf{s}^- \geq \mathbf{0}, \\ & Y\lambda - \mathbf{s}^+ \geq \mathbf{y}_o, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{6}$$

where $\theta = 1 - \beta$. Program (6) corresponds to the input-oriented CRS radial model, also known as CCR model. So, according to Remark 2.1 and considering strictly positive weights, the concept of efficiency given in Definition 2.1 coincides with the classical *CCR-efficiency* [27, Chapter 3].

On the other hand, if $D^- = 0$ (i.e. $d_1^-, \dots, d_m^- = 0$), and $D^+ = I_s$ (i.e. $d_1^+, \dots, d_s^+ = 1$), then we can write (4) as

$$\begin{aligned} \max_{\phi,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \phi + \varepsilon \left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & X\lambda + \mathbf{s}^- \leq \mathbf{x}_o, \\ & \phi\mathbf{y}_o - Y\lambda + \mathbf{s}^+ \leq \mathbf{0}, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{7}$$

where $\phi = 1 + \beta$. Program (7) is the output-oriented version of the CCR model.

Moreover, if $D^- = I_m$ and $D^+ = I_s$ (i.e. $d_1^-, \dots, d_m^-, d_1^+, \dots, d_s^+ = 1$), then model (4) is non-oriented and β^* coincides with the *generalized Farrell measure* [12].

3. Chance constrained directional models

We consider $D = \{\text{DMU}_1, \dots, \text{DMU}_n\}$ a set of n DMUs with m stochastic inputs and s stochastic outputs. Matrices $\tilde{X} = (\tilde{x}_{ij})$ and $\tilde{Y} = (\tilde{y}_{rj})$ are the *input* and *output data matrices*, respectively, where $\tilde{x}_{ij}$ and $\tilde{y}_{rj}$ represent the ith input and rth output of the jth DMU. Moreover, we denote by $X = (x_{ij})$ and $Y = (y_{rj})$ their expected values, which we assume to be strictly positive. Given $\text{DMU}_o \in D$, we define the column vectors $\tilde{\mathbf{x}}_o = (\tilde{x}_{1o}, \dots, \tilde{x}_{mo})^\top$, $\tilde{\mathbf{y}}_o = (\tilde{y}_{1o}, \dots, \tilde{y}_{so})^\top$, $\mathbf{x}_o = (x_{1o}, \dots, x_{mo})^\top$ and $\mathbf{y}_o = (y_{1o}, \dots, y_{so})^\top$. Although we suppose CRS in all programs, different returns to scale can be easily considered by adding the corresponding constraints: $\mathbf{e}\lambda = 1$ (VRS), $0 \leq \mathbf{e}\lambda \leq 1$ (NIRS), $\mathbf{e}\lambda \geq 1$ (NDRS) or $L \leq \mathbf{e}\lambda \leq U$ (GRS), with $0 \leq L \leq 1$ and $U \geq 1$.

In general, there are two methods for constructing a chance constrained model from a deterministic model in which inputs and outputs become stochastic, according to [28, Chapter 4]. On the one hand, a *P-model* is based on the highest probability of occurrence of the objective function. On the other hand, an *E-model* is based on the mathematical expectation to obtain the expected value of the objective function. In both cases, constraints must be satisfied with a probability greater than a certain threshold.

In particular, we want to adapt the deterministic directional models (3) or (4) to their chance constrained versions in which inputs and outputs become stochastic. Since the objective function remains deterministic, we are going to apply the E-model approach.

Note that, although programs (3) and (4) are equivalent, there is a relevant difference if we want to adapt them to the stochastic case: the directions $\mathbf{g}^-$ and $\mathbf{g}^+$ from (3) remain deterministic but, on the other hand, the directions $\mathbf{g}^- = D^-\tilde{\mathbf{x}}_o$ and $\mathbf{g}^+ = D^+\tilde{\mathbf{y}}_o$ from (4) become stochastic because they depend on the (stochastic) inputs and outputs. So, we are mostly interested in adapting program (4) because it generalizes the radial models, in which the radial direction depends on the inputs or outputs. Hence, we will first adapt (4) in Section 3.1 and later, adapt (3) in Section 3.2.

With respect to the nature of the directions, we use stochastic directions if the improvement strategy is relative to the value of the variables and, on the other hand, we use deterministic directions if the improvement strategy can be expressed in absolute terms, regardless of the value of the variables. For example, if the improvement strategy is increasing all outputs by the same percentage, the output direction is stochastic and the same for all outputs; on the other hand, if the improvement strategy is increasing all outputs by the same absolute quantity, the output direction is deterministic and the same for all outputs. We are going to show differences between stochastic and deterministic directions using an example in Section 4.

3.1. Stochastic directions

In this section, we are going to apply the E-model approach to program (4), considering that inputs and outputs become stochastic. Hence, given $0 < \alpha < 1$, the corresponding chance constrained E-model can be written as

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon \left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & P\left\{\left(\Theta^-(\beta)\tilde{\mathbf{x}}_o - \tilde{X}\lambda - \mathbf{s}^-\right)_i \geq 0\right\} \geq 1 - \alpha, \quad i = 1, \dots, m, \\ & P\left\{\left(\Theta^+(\beta)\tilde{\mathbf{y}}_o - \tilde{Y}\lambda + \mathbf{s}^+\right)_r \leq 0\right\} \geq 1 - \alpha, \quad r = 1, \dots, s, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{8}$$

where P denotes the probability function and $\Theta^-(\beta)$, $\Theta^+(\beta)$ are given by (5).

Definition 3.1. $\text{DMU}_o \in D$ is *α-stochastically efficient* if and only if the optimal objective value of (8) is 0, i.e. $\beta^* = 0$ and $\mathbf{w}^-\mathbf{s}^{-*} + \mathbf{w}^+\mathbf{s}^{+*} = 0$. Moreover, we say that DMU_o is *α-stochastically weakly efficient* if $\beta^* = 0$ and it is not α-stochastically efficient.

Remark 3.1. The concept of stochastic efficiency given in Definition 3.1 does not depend on the orientation $(\mathbf{d}^-, \mathbf{d}^+)$ because $\beta^* = 0$ and hence, by (5), $\Theta^-(\beta^*) = I_m$ and $\Theta^+(\beta^*) = I_s$ do not depend on the orientation. So, according to (8), optimal solutions remain optimal if we change the orientation, concluding that if DMU_o is α-stochastically efficient for a given orientation, then it is α-stochastically efficient for any orientation.

For constraints in (8), we have

$$P\left\{\left(\Theta^-(\beta)\tilde{\mathbf{x}}_o - \tilde{X}\lambda\right)_i \leq s_i^-\right\} \leq \alpha, \quad i = 1, \dots, m, \tag{9}$$

$$P\left\{\left(\tilde{Y}\lambda - \Theta^+(\beta)\tilde{\mathbf{y}}_o\right)_r \leq s_r^+\right\} \leq \alpha, \quad r = 1, \dots, s. \tag{10}$$

Let us assume that inputs and outputs are correlated random variables following multivariate normal distributions with means $E(\tilde{x}_{ij}) = x_{ij}$ and $E(\tilde{y}_{rj}) = y_{rj}$. The use of the normal distribution is due to its great versatility and it is discussed in [19]. Then, we can define standard

normal random variables by

$$\tilde{Z}_i^- = \frac{\left(\Theta^-(\beta)\tilde{\mathbf{x}}_o - \tilde{X}\lambda\right)_i - \left(\Theta^-(\beta)\mathbf{x}_o - X\lambda\right)_i}{\sigma_i^-(\beta,\lambda)}, \qquad i = 1,\dots,m,$$

$$\tilde{Z}_r^+ = \frac{\left(\tilde{Y}\lambda - \Theta^+(\beta)\tilde{\mathbf{y}}_o\right)_r - \left(Y\lambda - \Theta^+(\beta)\mathbf{y}_o\right)_r}{\sigma_r^+(\beta,\lambda)}, \qquad r = 1,\dots,s,$$

where

$$\begin{aligned}\left(\sigma_i^-(\beta,\lambda)\right)^2 &= \mathrm{Var}\left(\Theta^-(\beta)\tilde{\mathbf{x}}_o - \tilde{X}\lambda\right)_i = \mathrm{Var}\left((1-\beta d_i^-)\tilde{x}_{io} - \sum_{j=1}^n \lambda_j \tilde{x}_{ij}\right)\\ &= \sum_{j,q=1}^n \lambda_j\lambda_q \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{iq}) - 2(1-\beta d_i^-)\sum_{j=1}^n \lambda_j \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{io})\\ &\quad + (1-\beta d_i^-)^2 \mathrm{Var}(\tilde{x}_{io}), \quad i = 1,\dots,m,\end{aligned} \tag{11}$$

$$\begin{aligned}\left(\sigma_r^+(\beta,\lambda)\right)^2 &= \mathrm{Var}\left(\tilde{Y}\lambda - \Theta^+(\beta)\tilde{\mathbf{y}}_o\right)_r = \mathrm{Var}\left(\sum_{j=1}^n \lambda_j \tilde{y}_{rj} - (1+\beta d_r^+)\tilde{y}_{ro}\right)\\ &= \sum_{j,q=1}^n \lambda_j\lambda_q \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{rq}) - 2(1+\beta d_r^+)\sum_{j=1}^n \lambda_j \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{ro})\\ &\quad + (1+\beta d_r^+)^2 \mathrm{Var}(\tilde{y}_{ro}), \quad r = 1,\dots,s.\end{aligned} \tag{12}$$

Hence, (9) and (10) become

$$P\left\{\tilde{Z}_i^- \le \frac{s_i^- - \left(\Theta^-(\beta)\mathbf{x}_o - X\lambda\right)_i}{\sigma_i^-(\beta,\lambda)}\right\} \le \alpha, \quad i = 1,\dots,m, \tag{13}$$

$$P\left\{\tilde{Z}_r^+ \le \frac{s_r^+ - \left(Y\lambda - \Theta^+(\beta)\mathbf{y}_o\right)_r}{\sigma_r^+(\beta,\lambda)}\right\} \le \alpha, \quad r = 1,\dots,s. \tag{14}$$

Since $\tilde{Z}_i^-$ and $\tilde{Z}_r^+$ follow the standard normal distribution Φ, we can write (13) and (14) as

$$\frac{s_i^- - \left(\Theta^-(\beta)\mathbf{x}_o - X\lambda\right)_i}{\sigma_i^-(\beta,\lambda)} \le \Phi^{-1}(\alpha), \quad i = 1,\dots,m, \tag{15}$$

$$\frac{s_r^+ - \left(Y\lambda - \Theta^+(\beta)\mathbf{y}_o\right)_r}{\sigma_r^+(\beta,\lambda)} \le \Phi^{-1}(\alpha), \quad r = 1,\dots,s. \tag{16}$$

Reordering and vectorizing expressions (15) and (16), we have that the deterministic equivalent to (8) is given by

$$\begin{aligned}\max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right)\\ \text{s.t.} \quad & \Theta^-(\beta)\mathbf{x}_o - X\lambda - \mathbf{s}^- + \Phi^{-1}(\alpha)\boldsymbol{\sigma}^-(\beta,\lambda) \ge \mathbf{0},\\ & \Theta^+(\beta)\mathbf{y}_o - Y\lambda + \mathbf{s}^+ - \Phi^{-1}(\alpha)\boldsymbol{\sigma}^+(\beta,\lambda) \le \mathbf{0},\\ & \lambda \ge \mathbf{0},\ \mathbf{s}^- \ge \mathbf{0},\ \mathbf{s}^+ \ge \mathbf{0},\end{aligned} \tag{17}$$

that is equivalent to

$$\begin{aligned}\max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right)\\ \text{s.t.} \quad & \Theta^-(\beta)\mathbf{x}_o - X\lambda - \mathbf{s}^- + \Phi^{-1}(\alpha)\boldsymbol{\sigma}^-(\beta,\lambda) = \mathbf{0},\\ & \Theta^+(\beta)\mathbf{y}_o - Y\lambda + \mathbf{s}^+ - \Phi^{-1}(\alpha)\boldsymbol{\sigma}^+(\beta,\lambda) = \mathbf{0},\\ & \lambda \ge \mathbf{0},\ \mathbf{s}^- \ge \mathbf{0},\ \mathbf{s}^+ \ge \mathbf{0},\end{aligned} \tag{18}$$

because any optimal solution of (17) should also satisfy the constraints of (18), and vice versa. Hence, by Definition 3.1, DMU_o is α-stochastically efficient if and only if the optimal objective value of (17)–(18) is 0. Note that, considering $\boldsymbol{\sigma}^-, \boldsymbol{\sigma}^+$ as variables and adding (11), (12) as constraints, programs (17)–(18) are quadratically constrained. Moreover, for $0 < \alpha \le 0.5$, program (17) has a convex feasible set. In fact, it can be expressed as a second-order cone program.

As a particular case of (17), if $D^- = I_m$ (i.e. $d_1^-,\dots,d_m^- = 1$), and $D^+ = 0$ (i.e. $d_1^+,\dots,d_s^+ = 0$), then we can write

$$\begin{aligned}\min_{\theta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \theta - \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right)\\ \text{s.t.} \quad & \theta\mathbf{x}_o - X\lambda - \mathbf{s}^- + \Phi^{-1}(\alpha)\boldsymbol{\sigma}^-(\theta,\lambda) \ge \mathbf{0},\\ & Y\lambda - \mathbf{s}^+ + \Phi^{-1}(\alpha)\boldsymbol{\sigma}^+(\lambda) \ge \mathbf{y}_o,\\ & \lambda \ge \mathbf{0},\ \mathbf{s}^- \ge \mathbf{0},\ \mathbf{s}^+ \ge \mathbf{0},\end{aligned} \tag{19}$$

where $\theta = 1-\beta$, and

$$\begin{aligned}\left(\sigma_i^-(\theta,\lambda)\right)^2 &= \sum_{j,q=1}^n \lambda_j\lambda_q \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{iq}) - 2\theta\sum_{j=1}^n \lambda_j \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{io}) + \theta^2\mathrm{Var}(\tilde{x}_{io}),\\ &\quad i = 1,\dots,m,\\ \left(\sigma_r^+(\lambda)\right)^2 &= \sum_{j,q=1}^n \lambda_j\lambda_q \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{rq}) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{ro}) + \mathrm{Var}(\tilde{y}_{ro}),\\ &\quad r = 1,\dots,s.\end{aligned}$$

Program (19) is the chance constrained E-model corresponding to the input-oriented CCR model (6), firstly introduced with some simplifications by [17].

On the other hand, if $D^- = 0$ (i.e. $d_1^-,\dots,d_m^- = 0$), and $D^+ = I_s$ (i.e. $d_1^+,\dots,d_s^+ = 1$), then we can write (17) as

$$\begin{aligned}\max_{\phi,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \phi + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right)\\ \text{s.t.} \quad & X\lambda + \mathbf{s}^- - \Phi^{-1}(\alpha)\boldsymbol{\sigma}^-(\lambda) \le \mathbf{x}_o,\\ & \phi\mathbf{y}_o - Y\lambda + \mathbf{s}^+ - \Phi^{-1}(\alpha)\boldsymbol{\sigma}^+(\phi,\lambda) \le \mathbf{0},\\ & \lambda \ge \mathbf{0},\ \mathbf{s}^- \ge \mathbf{0},\ \mathbf{s}^+ \ge \mathbf{0},\end{aligned} \tag{20}$$

where $\phi = 1+\beta$, and

$$\begin{aligned}\left(\sigma_i^-(\lambda)\right)^2 &= \sum_{j,q=1}^n \lambda_j\lambda_q \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{iq}) - 2\sum_{j=1}^n \lambda_j \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{io}) + \mathrm{Var}(\tilde{x}_{io}),\\ &\quad i = 1,\dots,m,\\ \left(\sigma_r^+(\phi,\lambda)\right)^2 &= \sum_{j,q=1}^n \lambda_j\lambda_q \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{rq}) - 2\phi\sum_{j=1}^n \lambda_j \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{ro}) + \phi^2\mathrm{Var}(\tilde{y}_{ro}),\\ &\quad r = 1,\dots,s.\end{aligned}$$

Program (20) is the chance constrained E-model corresponding to the output-oriented CCR model (7). In fact, model (20) (with unit weights) was constructed in [21].

Remark 3.2. Both chance constrained radial models, the input-oriented (19) and the output-oriented (20), were used in [17,21], respectively, to define concepts of stochastic efficiency. Nevertheless, taking into account Remark 3.1 and the fact that (19) and (20) are particular cases of (17)–(18) (that are equivalent to (8)) with $\theta = 1-\beta$ and $\phi = 1+\beta$, we obtain that the concept of stochastic efficiency given in Definition 3.1 coincides with those introduced in [17,21].

Finally, if $D^- = I_m$ and $D^+ = I_s$ (i.e. $d_1^-,\dots,d_m^-,d_1^+,\dots,d_s^+ = 1$), then we obtain a chance constrained E-model for computing a *stochastic generalized Farrell measure* as the optimal value β^* of the following quadratically constrained program:

$$\begin{aligned}\max_{\beta,\lambda} \quad & \beta\\ \text{s.t.} \quad & (1-\beta)I_m\mathbf{x}_o - X\lambda + \Phi^{-1}(\alpha)\boldsymbol{\sigma}^-(\beta,\lambda) \ge \mathbf{0},\\ & (1+\beta)I_s\mathbf{y}_o - Y\lambda - \Phi^{-1}(\alpha)\boldsymbol{\sigma}^+(\beta,\lambda) \le \mathbf{0},\\ & \lambda \ge \mathbf{0},\end{aligned} \tag{21}$$

where $\boldsymbol{\sigma}^-(\beta,\lambda)$ and $\boldsymbol{\sigma}^+(\beta,\lambda)$ are given by (11) and (12), respectively. Moreover, for $0 < \alpha \le 0.5$, program (21) has a convex feasible set and can be expressed as a second-order cone program.

3.2. Deterministic directions

In this section, we are going to apply the E-model approach to program (3), considering that inputs and outputs become stochastic but the directions $\mathbf{g}^-$, $\mathbf{g}^+$ remain deterministic. Hence, given $0 < \alpha < 1$, the corresponding chance constrained E-model can be written as

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & P\left\{\left(\tilde{\mathbf{x}}_o - \beta\mathbf{g}^- - \tilde{X}\lambda - \mathbf{s}^-\right)_i \geq 0\right\} \geq 1-\alpha, \quad i = 1,\dots,m, \\ & P\left\{\left(\tilde{\mathbf{y}}_o + \beta\mathbf{g}^+ - \tilde{Y}\lambda + \mathbf{s}^+\right)_r \leq 0\right\} \geq 1-\alpha, \quad r = 1,\dots,s, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}. \end{aligned} \tag{22}$$

Remark 3.3. The optimal objective value of (22) is 0 if and only if the optimal objective value of (8) is 0, regardless of the directions. Hence, if we use program (22) instead of program (8) in Definition 3.1, then we obtain an equivalent concept of stochastic efficiency. Analogously, the same applies to the concept of stochastic weak efficiency.

If we assume that inputs and outputs are correlated random variables following multivariate normal distributions with means $E(\tilde{x}_{ij}) = x_{ij}$ and $E(\tilde{y}_{rj}) = y_{rj}$, we can deduce the deterministic equivalent to (22) following the methodology given in Section 3.1:

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & \beta\mathbf{g}^- + X\lambda + \mathbf{s}^- - \boldsymbol{\Phi}^{-1}(\alpha)\boldsymbol{\sigma}^-(\lambda) \leq \mathbf{x}_o, \\ & -\beta\mathbf{g}^+ + Y\lambda - \mathbf{s}^+ + \boldsymbol{\Phi}^{-1}(\alpha)\boldsymbol{\sigma}^+(\lambda) \geq \mathbf{y}_o, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{23}$$

where

$$\left(\sigma_i^-(\lambda)\right)^2 = \sum_{j,q=1}^{n} \lambda_j\lambda_q \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{iq}) - 2\sum_{j=1}^{n}\lambda_j \mathrm{Cov}(\tilde{x}_{ij},\tilde{x}_{io}) + \mathrm{Var}(\tilde{x}_{io}), \quad i = 1,\dots,m, \tag{24}$$

$$\left(\sigma_r^+(\lambda)\right)^2 = \sum_{j,q=1}^{n} \lambda_j\lambda_q \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{rq}) - 2\sum_{j=1}^{n}\lambda_j \mathrm{Cov}(\tilde{y}_{rj},\tilde{y}_{ro}) + \mathrm{Var}(\tilde{y}_{ro}), \quad r = 1,\dots,s. \tag{25}$$

Considering $\boldsymbol{\sigma}^-, \boldsymbol{\sigma}^+$ as variables and adding (24), (25) as constraints, program (23) is convex for $0 < \alpha \leq 0.5$ and it can be expressed as a second-order cone program. Although model (23) does not generalize chance constrained radial models due to the deterministic behavior of its directions, it may be appropriate in certain cases, as we will show in Section 4.

Remark 3.4. Given the model (17) with stochastic directions of the form $\mathbf{g}^- = D^-\tilde{\mathbf{x}}_o$ and $\mathbf{g}^+ = D^+\tilde{\mathbf{y}}_o$, we can consider a model (23) with deterministic directions given by the means of the stochastic directions, i.e. $\mathbf{g}^- = D^-\mathbf{x}_o$ and $\mathbf{g}^+ = D^+\mathbf{y}_o$. We can write this model as:

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & \Theta^-(\beta)\mathbf{x}_o - X\lambda - \mathbf{s}^- + \boldsymbol{\Phi}^{-1}(\alpha)\boldsymbol{\sigma}^-(\lambda) \geq \mathbf{0}, \\ & \Theta^+(\beta)\mathbf{y}_o - Y\lambda + \mathbf{s}^+ - \boldsymbol{\Phi}^{-1}(\alpha)\boldsymbol{\sigma}^+(\lambda) \leq \mathbf{0}, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{26}$$

where $\boldsymbol{\sigma}^+(\lambda)$ and $\boldsymbol{\sigma}^-(\lambda)$ are given by (24) and (25), respectively. In this case, we say that both models, (17) and (26), are *associated*. We have that associated models are equivalent in the deterministic case, i.e. if the variances of the stochastic variables are 0.

Conversely, given the model (23) with deterministic directions $\mathbf{g}^-$ and $\mathbf{g}^+$, we can always find an associated model (17) with stochastic directions of the form $\mathbf{g}^- = D^-\tilde{\mathbf{x}}_o$ and $\mathbf{g}^+ = D^+\tilde{\mathbf{y}}_o$, where $D^- = \mathrm{diag}(d_1^-,\dots,d_m^-)$ and $D^+ = \mathrm{diag}(d_1^+,\dots,d_s^+)$ with $d_i^- = g_i^-/x_{io}$, $d_r^+ = g_r^+/y_{ro}$.

3.3. Joint chance constrained directional models

The chance constrained methodology ensures that the constraints are met with a certain probability, but considering each constraint separately. A more appropriate study would be to require that all the constraints are met at once with a certain probability.

Thus, for stochastic directions, program (8) would become

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & P\left\{\begin{array}{l}\left(\Theta^-(\beta)\tilde{\mathbf{x}}_o - \tilde{X}\lambda - \mathbf{s}^- \geq \mathbf{0}\right) \wedge \\ \left(\Theta^+(\beta)\tilde{\mathbf{y}}_o - \tilde{Y}\lambda + \mathbf{s}^+ \leq \mathbf{0}\right)\end{array}\right\} \geq 1-\alpha, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{27}$$

and for deterministic directions, program (22) would become

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & P\left\{\begin{array}{l}\left(\tilde{\mathbf{x}}_o - \beta\mathbf{g}^- - \tilde{X}\lambda - \mathbf{s}^- \geq \mathbf{0}\right) \wedge \\ \left(\tilde{\mathbf{y}}_o + \beta\mathbf{g}^+ - \tilde{Y}\lambda + \mathbf{s}^+ \leq \mathbf{0}\right)\end{array}\right\} \geq 1-\alpha, \\ & \lambda \geq \mathbf{0},\ \mathbf{s}^- \geq \mathbf{0},\ \mathbf{s}^+ \geq \mathbf{0}, \end{aligned} \tag{28}$$

where $\wedge$ denotes the logical conjunction. But, in order to find the deterministic equivalent program of (27) or (28), we should use multivariate probability distributions, for which the inverse distribution functions are not uniquely defined.

An approach to overcome this problem is given in [29], whose authors consider stochastically independent variables, so that the joint probability is the product of the separate probabilities. Nevertheless, this assumption reduces the applicability of the model and, moreover, it leads to highly non-linear programs that are very hard to solve, as we will show below. Hence, assuming that the variables are stochastically independent and following the methodology in [29], program (27) can be written as

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+,\boldsymbol{\gamma}^-,\boldsymbol{\gamma}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & P\left\{\left(\Theta^-(\beta)\tilde{\mathbf{x}}_o - \tilde{X}\lambda\right)_i \leq s_i^-\right\} \leq \alpha^{\gamma_i^-}, \quad i = 1,\dots,m, \\ & P\left\{\left(\tilde{Y}\lambda - \Theta^+(\beta)\tilde{\mathbf{y}}_o\right)_r \leq s_r^+\right\} \leq \alpha^{\gamma_r^+}, \quad r = 1,\dots,s, \\ & \mathbf{e}^-\boldsymbol{\gamma}^- = 1, \quad \mathbf{e}^+\boldsymbol{\gamma}^+ = 1, \\ & \lambda \geq \mathbf{0}, \quad \mathbf{s}^-,\boldsymbol{\gamma}^- \geq \mathbf{0}, \quad \mathbf{s}^+,\boldsymbol{\gamma}^+ \geq \mathbf{0}, \end{aligned} \tag{29}$$

for stochastic directions, where $\boldsymbol{\gamma}^- \in \mathbb{R}^m$, $\boldsymbol{\gamma}^+ \in \mathbb{R}^s$ are auxiliary parameters and $\mathbf{e}^- \in \mathbb{R}^m$, $\mathbf{e}^+ \in \mathbb{R}^s$ are the all-ones row vectors. The corresponding deterministic equivalent is given by

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+,\boldsymbol{\gamma}^-,\boldsymbol{\gamma}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & \Theta^-(\beta)\mathbf{x}_o - X\lambda - \mathbf{s}^- + \boldsymbol{\Phi}^{-1}(\alpha^{\boldsymbol{\gamma}^-})\boldsymbol{\sigma}^-(\beta,\lambda) \geq \mathbf{0}, \\ & \Theta^+(\beta)\mathbf{y}_o - Y\lambda + \mathbf{s}^+ - \boldsymbol{\Phi}^{-1}(\alpha^{\boldsymbol{\gamma}^+})\boldsymbol{\sigma}^+(\beta,\lambda) \leq \mathbf{0}, \\ & \mathbf{e}^-\boldsymbol{\gamma}^- = 1, \quad \mathbf{e}^+\boldsymbol{\gamma}^+ = 1, \\ & \lambda \geq \mathbf{0}, \quad \mathbf{s}^-,\boldsymbol{\gamma}^- \geq \mathbf{0}, \quad \mathbf{s}^+,\boldsymbol{\gamma}^+ \geq \mathbf{0}, \end{aligned} \tag{30}$$

where $\boldsymbol{\Phi}^{-1}(\alpha^{\boldsymbol{\gamma}^-}) \in \mathbb{R}^m$ and $\boldsymbol{\Phi}^{-1}(\alpha^{\boldsymbol{\gamma}^+}) \in \mathbb{R}^s$ are vectors with components $\boldsymbol{\Phi}^{-1}(\alpha^{\gamma_i^-})$ for $i = 1,\dots,m$ and $\boldsymbol{\Phi}^{-1}(\alpha^{\gamma_r^+})$ for $r = 1,\dots,s$, respectively. Moreover, $\boldsymbol{\sigma}^-(\beta,\lambda)$ and $\boldsymbol{\sigma}^+(\beta,\lambda)$ are given by (11) and (12), respectively. Considering $\boldsymbol{\sigma}^-, \boldsymbol{\sigma}^+$ as variables and adding (11), (12) as constraints, program (30) is convex for $0 < \alpha \leq 0.5$ and it can be expressed as a second-order cone program.

Table 1
Average values from the first 10 school sites in the Program Follow Through.

	Inputs	Outputs
Site 1	(86.13, 16.24, 48.21, 49.69, 9)	(54.53, 58.98, 38.16)
Site 2	(29.26, 10.24, 41.96, 40.65, 5)	(24.69, 33.89, 26.02)
Site 3	(43.12, 11.31, 38.19, 35.03, 9)	(36.41, 40.62, 28.51)
Site 4	(24.96, 6.14, 24.81, 25.15, 7)	(14.94, 17.58, 16.19)
Site 5	(11.62, 2.21, 6.85, 6.37, 4)	(7.81, 6.94, 5.37)
Site 6	(11.88, 4.97, 18.73, 18.04, 4)	(12.59, 16.85, 12.84)
Site 7	(32.64, 6.88, 28.10, 25.45, 7)	(17.06, 16.99, 17.82)
Site 8	(20.79, 12.97, 54.85, 52.07, 8)	(20.19, 30.64, 33.16)
Site 9	(34.40, 11.04, 38.16, 42.40, 8)	(26.13, 29.80, 26.29)
Site 10	(61.74, 14.50, 49.09, 42.92, 9)	(46.42, 51.59, 35.20)

Analogously, for deterministic directions, program (28) can be written as

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+,\boldsymbol{\gamma}^-,\boldsymbol{\gamma}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & P\left\{\left(\tilde{\mathbf{x}}_o - \beta\mathbf{g}^- - \tilde{X}\lambda\right)_i \le s_i^-\right\} \le \alpha^{\gamma_i^-}, \quad i = 1,\dots,m, \\ & P\left\{\left(\tilde{Y}\lambda - \tilde{\mathbf{y}}_o - \beta\mathbf{g}^+\right)_r \le s_r^+\right\} \le \alpha^{\gamma_r^+}, \quad r = 1,\dots,s, \\ & \mathbf{e}^-\boldsymbol{\gamma}^- = 1, \quad \mathbf{e}^+\boldsymbol{\gamma}^+ = 1, \\ & \lambda \ge \mathbf{0}, \quad \mathbf{s}^-,\boldsymbol{\gamma}^- \ge \mathbf{0}, \quad \mathbf{s}^+,\boldsymbol{\gamma}^+ \ge \mathbf{0}, \end{aligned} \tag{31}$$

whose deterministic equivalent is given by

$$\begin{aligned} \max_{\beta,\lambda,\mathbf{s}^-,\mathbf{s}^+,\boldsymbol{\gamma}^-,\boldsymbol{\gamma}^+} \quad & \beta + \varepsilon\left(\mathbf{w}^-\mathbf{s}^- + \mathbf{w}^+\mathbf{s}^+\right) \\ \text{s.t.} \quad & \beta\mathbf{g}^- + X\lambda + \mathbf{s}^- - \boldsymbol{\Phi}^{-1}(\alpha^{\boldsymbol{\gamma}^-})\boldsymbol{\sigma}^-(\lambda) \le \mathbf{x}_o, \\ & -\beta\mathbf{g}^+ + Y\lambda - \mathbf{s}^+ + \boldsymbol{\Phi}^{-1}(\alpha^{\boldsymbol{\gamma}^+})\boldsymbol{\sigma}^+(\lambda) \ge \mathbf{y}_o, \\ & \mathbf{e}^-\boldsymbol{\gamma}^- = 1, \quad \mathbf{e}^+\boldsymbol{\gamma}^+ = 1, \\ & \lambda \ge \mathbf{0}, \quad \mathbf{s}^-,\boldsymbol{\gamma}^- \ge \mathbf{0}, \quad \mathbf{s}^+,\boldsymbol{\gamma}^+ \ge \mathbf{0}, \end{aligned} \tag{32}$$

where $\boldsymbol{\sigma}^-(\lambda)$ and $\boldsymbol{\sigma}^+(\lambda)$ are given by (24) and (25), respectively. Considering $\boldsymbol{\sigma}^-, \boldsymbol{\sigma}^+$ as variables and adding (24), (25) as constraints, program (31) is convex for $0 < \alpha \le 0.5$ and it can be expressed as a second-order cone program.

As we noted above, programs (30) and (32) are highly non-linear due to the new unknown parameters $\boldsymbol{\gamma}^- \in \mathbb{R}^m$ and $\boldsymbol{\gamma}^+ \in \mathbb{R}^s$. However, some approximation methods based on second order cone programming can be utilized as it is shown in [29]. This certainly poses new open problems for future research.

4. Examples

We are going to consider the example given in [3,17], in which the "Program Follow Through" experiment is analyzed in public school education. We have to note that this example is given for illustrative purposes only. Some other meaningful examples could be given by considering factories instead of schools, workers instead of students, etc. The original data set consists on 49 school sites enrolled in the experiment, with 5 inputs (education, level of mother, parent occupation, parental visit index, counseling index, and number of teachers) and 3 outputs (total reading scores, total math scores, and total Coopersmith[1] scores). The average values of input and output variables are reported for each school site. Table 1 shows these values for the first 10 school sites.

In order to simplify the problem, we are going to assume that inputs are deterministic and outputs are stochastic. Moreover, we assume that all observed outputs coincide with their mathematical expectations, all outputs are stochastically independent, and the within-school variability of each output (as measured by the variance) is the same, c^2, for all outputs and at all school sites. See [17] for a detailed discussion on these assumptions.

In this example, we are interested in knowing how much each school site's output scores must improve to become α-stochastically efficient or weakly efficient and therefore, only the optimal score β^* of an output-oriented model (i.e. with $\mathbf{g}^- = \mathbf{0}$ or $D^- = 0$) will be necessary. The rationale for considering weak efficiency as a goal (jointly with efficiency) lies in the fact that any small improvement in some output that has no slacks, results in the DMU being α-stochastically efficient. Note that this is always possible since, in output-oriented models, there is always some output without slacks.

If we take $D^- = 0$ and $D^+ = I_3$ (i.e. $d_1^+ = d_2^+ = d_3^+ = 1$) in program (17), then the chance constrained output-oriented CCR model (20) is applied. However, this choice treats all evaluated subjects (reading, math and Coopersmith) in the same way and hence, it may not reflect the specific characteristics of improvement of each subject. On the other hand, if for example we set $d_1^+ = 0.1$, $d_2^+ = 0.05$ and $d_3^+ = 0.01$, then we suppose that students' ability to improve the reading score by 10% is the same as their ability to improve the math score by 5%, and the same as their ability to improve the Coopersmith score by 1%.

If we look at improvement capabilities in absolute terms, then the directions become deterministic and we must apply model (23) with $\mathbf{g}^- = \mathbf{0}$. For example, if we set $g_1^+ = 5$, $g_2^+ = 4$ and $g_3^+ = 1$, then we suppose that students' ability to improve the reading score by 5 absolute points is the same as their ability to improve the math score by 4 absolute points, and the same as their ability to improve the Coopersmith score by 1 absolute point.

Note that, in order to compare optimal scores β^* from models with different directions, we have to take into account that those directions should represent the same "amount of effort" on the part of the students of the corresponding evaluated school site. For example, if the students in the Site 1 can improve their reading scores by 10%, their math scores by 5% and their Coopersmith scores by 1% using 1 "unit of effort", and the students in the Site 2 can improve their reading scores by 5 points, their math scores by 4 points and their Coopersmith scores by 1 point also using 1 "unit of effort", then we can compare the β^* scores of both sites if we use model (17) with $D^- = 0$, $d_1^+ = 0.1$, $d_2^+ = 0.05$, $d_3^+ = 0.01$ for evaluating Site 1, and model (23) with $\mathbf{g}^- = \mathbf{0}$, $g_1^+ = 5$, $g_2^+ = 4$, $g_3^+ = 1$ for evaluating Site 2. In this case, β^* is interpreted as the "total amount of effort" of the corresponding school site for becoming α-stochastically efficient or weakly efficient.

Taking all this into account, we have evaluated the first 10 school sites with respect to the whole sample, for $\alpha = 0.05$ and using different values c^2 of the variance of the outputs, obtaining the deterministic case by taking $c = 0$. Tables 2–4 show the results of the optimal β^* scores of chance constrained directional models with stochastic directions (columns 1 – 3) and their associated models with deterministic directions (columns 4 – 6). Specifically, Table 2 shows the results of applying the output-oriented chance constrained CCR model (20), that does not reflect the specific characteristics of improvement of each subject. Analogously, Table 3 shows the results of applying the chance constrained directional model (17) with stochastic directions determined by orientation parameters $D^- = 0$, $d_1^+ = 0.1$, $d_2^+ = 0.05$ and $d_3^+ = 0.01$. In this case, we are assuming that the students can improve the reading score by 10% the math score by 5% and the Coopersmith score by 1% using 1 "unit of effort" in all school sites. Finally, Table 4 shows the results of applying the chance constrained directional model (23) with deterministic directions given by $\mathbf{g}^- = \mathbf{0}$, $g_1^+ = 5$, $g_2^+ = 4$ and $g_3^+ = 1$, assuming that the students can improve the reading score by 5 points, the math score by 4 points and the Coopersmith score by 1 point using 1 "unit of effort" in all school sites.

Regarding the discussion of the results, first of all, associated models with stochastic and deterministic directions coincide in the deterministic case ($c = 0$), as stated in Remark 3.4. Moreover, associated models give similar results, as long as the variance of the stochastic variables

[1] An index of a child's self-esteem.

Table 2
Results of the optimal β^* scores of chance constrained directional models with stochastic directions ($D^- = 0$, $d_1^+ = 1$, $d_2^+ = 1$, $d_3^+ = 1$) and their associated models with deterministic directions ($\mathbf{g}^- = \mathbf{0}$, $g_1^+ = y_{1o}$, $g_2^+ = y_{2o}$, $g_3^+ = y_{3o}$), applied to the first 10 school sites using different values c^2 of the variance of the outputs and $\alpha = 0.05$.

Stochastic directions: $D^- = 0$, $d_1^+ = 1$, $d_2^+ = 1$, $d_3^+ = 1$										
Site	1	2	3	4	5	6	7	8	9	10
$c = 0$	0	0.109	0.012	0.108	0	0.103	0.121	0.093	0.148	0
$c = 0.5$	0	0.071	0	0.042	0	0.031	0.061	0.063	0.095	0
$c = 1$	0	0.036	0	0	0	0	0.006	0.026	0.053	0
Deterministic directions: $\mathbf{g}^- = \mathbf{0}$, $g_1^+ = y_{1o}$, $g_2^+ = y_{2o}$, $g_3^+ = y_{3o}$										
Site	1	2	3	4	5	6	7	8	9	10
$c = 0$	0	0.109	0.012	0.108	0	0.103	0.121	0.093	0.148	0
$c = 0.5$	0	0.073	0	0.044	0	0.033	0.063	0.065	0.098	0
$c = 1$	0	0.038	0	0	0	0	0.007	0.033	0.055	0

Table 3
Results of the optimal β^* scores of chance constrained directional models with stochastic directions ($D^- = 0$, $d_1^+ = 0.1$, $d_2^+ = 0.05$, $d_3^+ = 0.01$) and their associated models with deterministic directions ($\mathbf{g}^- = \mathbf{0}$, $g_1^+ = d_1^+ y_{1o}$, $g_2^+ = d_2^+ y_{2o}$, $g_3^+ = d_3^+ y_{3o}$), applied to the first 10 school sites using different values c^2 of the variance of the outputs and $\alpha = 0.05$.

Stochastic directions: $D^- = 0$, $d_1^+ = 0.1$, $d_2^+ = 0.05$, $d_3^+ = 0.01$										
Site	1	2	3	4	5	6	7	8	9	10
$c = 0$	0	5.041	0.388	4.988	0	3.380	5.468	8.218	5.303	0
$c = 0.5$	0	3.601	0	2.117	0	1.664	2.876	6.301	4.481	0
$c = 1$	0	2.296	0	0	0	0	0.374	3.409	3.573	0
Deterministic directions: $\mathbf{g}^- = \mathbf{0}$, $g_1^+ = d_1^+ y_{1o}$, $g_2^+ = d_2^+ y_{2o}$, $g_3^+ = d_3^+ y_{3o}$										
Site	1	2	3	4	5	6	7	8	9	10
$c = 0$	0	5.041	0.388	4.988	0	3.380	5.468	8.218	5.303	0
$c = 0.5$	0	3.707	0	2.216	0	1.768	2.994	6.437	4.592	0
$c = 1$	0	2.426	0	0	0	0	0.404	3.555	3.752	0

Table 4
Results of the optimal β^* scores of chance constrained directional models with deterministic directions ($\mathbf{g}^- = \mathbf{0}$, $g_1^+ = 5$, $g_2^+ = 4$, $g_3^+ = 1$) and their associated models with stochastic directions ($D^- = 0$, $d_1^+ = 5/y_{1o}$, $d_2^+ = 4/y_{2o}$, $d_3^+ = 1/y_{3o}$), applied to the first 10 school sites using different values c^2 of the variance of the outputs and $\alpha = 0.05$.

Stochastic directions: $D^- = 0$, $d_1^+ = 5/y_{1o}$, $d_2^+ = 4/y_{2o}$, $d_3^+ = 1/y_{3o}$										
Site	1	2	3	4	5	6	7	8	9	10
$c = 0$	0	1.982	0.211	1.137	0	0.754	1.412	3.090	2.561	0
$c = 0.5$	0	1.415	0	0.466	0	0.338	0.729	2.089	2.100	0
$c = 1$	0	0.819	0	0	0	0	0.080	1.130	1.413	0
Deterministic directions: $\mathbf{g}^- = \mathbf{0}$, $g_1^+ = 5$, $g_2^+ = 4$, $g_3^+ = 1$										
Site	1	2	3	4	5	6	7	8	9	10
$c = 0$	0	1.982	0.211	1.137	0	0.754	1.412	3.090	2.561	0
$c = 0.5$	0	1.457	0	0.487	0	0.359	0.755	2.134	2.152	0
$c = 1$	0	0.864	0	0	0	0	0.093	1.179	1.483	0

is not too large. Second, the greater the stochastic variability of outputs (the greater the coefficient c), the smaller the optimal score β^* is, and hence, school sites are less inefficient. The reason is that in the chance constrained formulation of DEA, the "hard" efficient frontier of deterministic DEA is replaced with a "soft" frontier that moves successively closer to any given observation [17]. This effect is also seen among associated models, where models with stochastic directions produce better scores β^* than the corresponding associated models with deterministic directions. Third, the relative rankings of inefficient school sites differ in the chance constrained analysis as compared to deterministic DEA. For example, in Tables 2 and 3, Site 7 is more inefficient than Site 2 for $c = 0$, but this situation is reversed for $c = 0.5$ and $c = 1$. Moreover, in Table 4, Site 8 is more inefficient than Site 9 for $c = 0$, but the opposite occurs for $c = 0.5$ and $c = 1$. Finally, different directions (representing different improvement strategies) also produce different rankings, as expected. For example, in Table 2, Site 2 is more inefficient than Site 8 for all values of c, just the opposite of what happens in Tables 3 and 4.

We have used R 4.2.0 [30] for computations. Specifically, we have used the optiSolve package [31] for solving quadratically constrained programs, and the deaR package [32] for linear models, in order to check the deterministic case $c = 0$.

5. Concluding remarks

We have constructed different versions of stochastic directional models following the chance constrained methodology and generalizing the existing radial models developed in [21]. The advantage of directional models over radial models is the great versatility provided by the choice of the direction, allowing the user to customize an improvement strategy that can be adapted to market conditions. Moreover, chance constrained models allow a closer approximation to a reality in which measurements present uncertainty, providing an alternative to fuzzy models. Hence, chance constrained directional models are appropriate in any situation with stochastic variables where the user wants to manage the proportion in which these variables are modified to achieve efficiency.

Nevertheless, chance constrained DEA also has drawbacks. The major operational disadvantage is that programs are quadratically constrained, as opposed to deterministic models whose programs are usually linear. However, the models presented in this work lie within the category of Second-Order Cone Programming, for which the numerical algorithm *Interior Point Optimization* is fast and robust, compared with general non-linear optimization methods. On the other hand, no formal statistical model with a sampling process is specified, unlike what is done in [16]. The main consequence is the difficulty in defining an appropriate and meaningful measure of inefficiency [15].

Finally, according to Section 3.3, a more appropriate chance constrained methodology would be to require that all the constraints are met at once with a certain probability, which would be a more general problem, broadening thus the scope of the applications of the models. However, in order to find the corresponding deterministic equivalent program, we should either consider independent variables (thus significantly reducing the applicability), or use multivariate probability distributions, for which the inverse distribution functions are not uniquely defined. Nevertheless, despite the difficulty of working with joint chance constrained models, they could lead to future studies and more sophisticated models in the chance constrained DEA framework.

CRediT authorship contribution statement

V.J. Bolós: Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **R. Benítez:** Writing – review & editing, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **V. Coll-Serrano:** Writing – review & editing, Validation, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

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

Data availability

Data will be made available on request.

References

[1] Charnes A, Cooper W, Rhodes E. Measuring the efficiency of decision making units. European J Oper Res 1978;2(6):429–44.

[2] Charnes A, Cooper W, Rhodes E. Short communication: Measuring the efficiency of decision making units. European J Oper Res 1979;3(4):339.

[3] Charnes A, Cooper WW, Rhodes E. Evaluating program and managerial efficiency: an application of data envelopment analysis to Program Follow Through. Manage Sci 1981;27(6):668–97.

[4] Banker RD, Charnes A, Cooper WW. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manage Sci 1984;30(9):1078–92.

[5] Färe R, Knox Lovell CA. Measuring the technical efficiency of production. J Econom Theory 1978;19(1):150–62.

[6] Charnes A, Cooper W, Golany B, Seiford L, Stutz J. Foundations of data envelopment analysis for Pareto-Koopmans efficient empirical production functions. J Econometrics 1985;30(1–2):91–107.

[7] Pastor J, Ruiz J, Sirvent I. An enhanced DEA Russell graph efficiency measure. European J Oper Res 1999;115(3):596–607.

[8] Tone K. A slacks-based measure of efficiency in data envelopment analysis. European J Oper Res 2001;130(3):498–509.

[9] Aparicio J, Ruiz JL, Sirvent I. Closest targets and minimum distance to the Pareto-efficient frontier in DEA. J Product Anal 2007;28(3):209–18.

[10] Tone K. Variations on the theme of slacks-based measure of efficiency in DEA. European J Oper Res 2010;200(3):901–7.

[11] Luenberger DG. New optimality principles for economic efficiency and equilibrium. J Optim Theory Appl 1992;75(2):221–64.

[12] Briec W. A graph-type extension of Farrell technical efficiency measure. J Product Anal 1997;8(1):95–110.

[13] Chambers RG, Chung Y, Färe R. Benefit and distance functions. J Econom Theory 1996;70(2):407–19.

[14] Chambers RG, Chung Y, Färe R. Profit, directional distance functions, and Nerlovian efficiency. J Optim Theory Appl 1998;98(2):351–64.

[15] Olesen OB, Petersen NC. Stochastic data envelopment analysis—A review. European J Oper Res 2016;251(1):2–21.

[16] Banker RD. Maximum likelihood, consistency and data envelopment analysis: A statistical foundation. Manage Sci 1993;39(10):1265–73.

[17] Land KC, Knox Lovell CA, Thore S. Chance-constrained data envelopment analysis. Manage Decis Econ 1993;14(6):541–54, https://www.jstor.org/stable/2487873.

[18] Olesen OB, Petersen NC. Chance constrained efficiency evaluation. Manage Sci 1995;41(3):442–57.

[19] Cooper WW, Huang ZM, Li SX. Satisficing DEA models under chance constraints. Ann Oper Res 1996;66:279–95.

[20] Cooper WW, Huang Z, Lelas V, Li SX, Olesen OB. Chance constrained programming formulations for stochastic characterizations of efficiency and dominance in DEA. J Product Anal 1998;9(1):53–79.

[21] Cooper WW, Deng H, Huang Z, Li SX. Chance constrained programming approaches to technical efficiencies and inefficiencies in stochastic data envelopment analysis. J Oper Res Soc 2002;53(12):1347–56.

[22] Xiao H, Liu X, Ren T, Zhou Z. Estimation of portfolio efficiency via stochastic DEA. RAIRO Oper Res 2022;56(4):2367–87.

[23] Kao C, Liu S-T. Fuzzy efficiency measures in data envelopment analysis. Fuzzy Sets and Systems 2000;113(3):427–37.

[24] Guo P, Tanaka H. Fuzzy DEA: A perceptual evaluation method. Fuzzy Sets and Systems 2001;119(1):149–60.

[25] León T, Liern V, Ruiz JL, Sirvent I. A fuzzy mathematical programming approach to the assessment of efficiency with DEA models. Fuzzy Sets and Systems 2003;139(2):407–19.

[26] Tavana M, Shiraz RK, Hatami-Marbini A, Agrell P, Paryab K. Chance-constrained DEA models with random fuzzy inputs and outputs. Knowl-Based Syst 2013;52:32–52.

[27] Cooper WW, Seiford LM, Tone K. Data envelopment analysis. A comprehensive text with models, applications, references and DEA-Solver software. 2nd ed. Springer; 2007, p. 483.

[28] Amirteimoori A, Sahoo BK, Charles V, Mehdizadeh S. Stochastic benchmarking. 1st ed. Springer Cham; 2022, p. 145.

[29] Shiraz RK, Tavana M, Fukuyama H. A joint chance-constrained data envelopment analysis model with random output data. Oper Res 2021;21:1255–77.

[30] Team RC. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2022.

[31] Wellmann R. optiSolve: Linear, quadratic, and rational optimization. 2021, R package version 1.0.

[32] Coll-Serrano V, Bolós VJ, Benítez R. deaR: Conventional and fuzzy data envelopment analysis. 2022, R package version 1.3.3.