Journal of Air Transport Management 58 (2017) 100–112

Contents lists available at ScienceDirect

Journal of Air Transport Management

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

Efficiency thresholds and cost structure in Senegal airports☆

Peter Wanke [a, *], C.P. Barros [b]

[a] COPPEAD Graduate Business School, Federal University of Rio de Janeiro, Rua Paschoal Lemme, 355, 21949-900, Rio de Janeiro, Brazil
[b] ISEG – Lisbon School of Economics and Management, ULisboa and CEsA - Research Centre on African, Asian and Latin American Studies, Rua Miguel Lupi, 20, 1249-078, Lisboa, Portugal

ARTICLE INFO

Article history:
Received 7 July 2016
Received in revised form
5 October 2016
Accepted 11 October 2016

Keywords:
Airports
Senegal
Satisficing DEA
Support vector machines
Performance thresholds
Cost structure

ABSTRACT

This paper assesses airport efficiency levels in Senegal under a stochastic environment where the satisficing concept for different performance thresholds is applied. A two-stage satisficing DEA-Support Vector Machine approach is used here to compute the impacts of cost structure on these thresholds. In the first stage, within the ambit of the satisficing DEA model, the probabilities of achieving a minimal performance threshold are computed in a stochastic fashion. In the second stage, Support Vector Machine regression is used to discriminate between high/low efficiency groups within a given performance threshold. This methodology was sufficiently robust to handle small samples. The results reveal that the cost of capital and the cost of labor are the cost structure variables that have the greatest impact on efficiency levels, besides cargo operations.

1. Introduction

This study focuses on the efficiency of Senegal's airports. Research into African airports is a recent and relatively understudied field in efficiency measurement (Wanke et al., 2016a). This may be due to certain factors, such as the poor quality of African data obtained from various economic sectors (Wanke et al., 2015a, 2015b) or reduced sample sizes (Barros and Wanke, 2015). Over the years, however, Data Envelopment Analysis (DEA) approaches have evolved to handle such shortcomings that may jeopardize the discriminatory power of efficiency scores, biasing them towards one.

One possible approach is to consider stochastic input and output sources in DEA by means of chance-constrained programming, developed by Charnes and Cooper (1963) and Kall (1976). In this approach, it is assumed that the efficiency of a DMU is stochastic, and the observation is an occurrence of a random phenomenon. Applications of chance-constrained DEA can be seen in studies that evaluate efficiency in some sectors around the world (Sueyoshi, 2000; Yang and Wen, 2005; Talluri et al., 2006; Li et al., 2007; Agpak and Gökçen, 2007; Yang et al., 2007; Chen, 2002; and Bhattacharya, 2009). To the best of our knowledge, however, no studies of airport efficiency have been performed using chance-constrained DEA.

It is worth mentioning that chance-constrained DEA approaches suffer from a major drawback in that they do not incorporate the concept of “satisficing”. The concept of “satisficing” has its origin in the psychology literature, where Simon (1957) used the term as an alternative to the assumption of “optimizing” behavior, which is extensively used in economics. Applications of satisficing DEA models are quite scarce. For a more recent contribution regarding satisficing DEA in the field of efficiency measurement, the reader is referred to the work of Tsolas and Charles (2015). In this research, a novel satisficing DEA model for measuring airport efficiency under a stochastic environment is presented in the case of Senegal. The proposed model is applied to Senegal's airport industry to assess probabilistically the efficiency of five airports during the 1996–2015 period - a twenty-year time span. By applying the bootstrapping technique for the generation of resampled inputs and outputs, it is possible to compute not only the efficiency probability distributions for each airport, but also their satisficing probabilities in terms of a given performance threshold (for example, what is the probability of the efficiency of a given airport being higher than 70%).

Despite the numerous studies focusing on airport efficiency and

☆ Research made with support of Calouste Gulbenkian Foundation.
* Corresponding author.
E-mail addresses: peter@coppead.ufrj.br, peter@coppead.ufrj.br (P. Wanke), cbarros@iseg.utl.pt (C.P. Barros).

http://dx.doi.org/10.1016/j.jairtraman.2016.10.005
0969-6997/© 2016 Elsevier Ltd. All rights reserved.

productivity using DEA and other stochastic frontier approaches, a satisficing DEA approach to the airport industry at the country level is still missing, thus suggesting a literature gap. As a matter of fact, the comprehensive literature review on airport efficiency presented in Damacena et al. (2016) and in Wanke et al. (2016a) indicates that most research on airport efficiency is conducted by means of non-parametric models such as DEA. It is only more recently that the stochastic element inherent to input/output measurement has been treated using the bootstrapping method, either in the form of data generating processes for the inputs and the outputs (Merkert and Pearson, 2015) or a bootstrapped truncated regression (Wanke et al., 2016a), but not in the form of chance-constrained programs subject to a probability distribution. Therefore, this research also adds to the current body of literature on airport efficiency, especially the part devoted to African countries, by focusing on another stream of possible approaches for treating the stochastic element in DEA.

As regards the lack of discriminatory power of efficiency scores when samples are small, Wanke et al. (2015a, 2015b, 2016b) showed the importance of using efficiency methods with high discriminatory power towards the efficiency frontier - i.e. lower efficiency scores in contrast to traditional DEA. The authors also advocate the combination of different predictive modeling techniques to explore effectively the impact of contextual variables on efficiency measurement in what is commonly known as two-stage DEA. This paper innovates in this context by adopting Support Vector Machine (SVM) regression in the second stage of analysis. SVM regression allows discrimination between high/low efficiency groups within each performance threshold in light of a given set of contextual variables, thus permitting the identification of the most significant efficiency drivers at each performance level.

The motivations for the present research are given next. Firstly, and justifying the present research, Senegal belongs to the region of Africa that is relatively unexplored in terms of airport efficiency. Secondly, this paper builds upon previous studies related to airport efficiency by evaluating the relative efficiency of Senegal's airports and their major drivers along a given performance threshold. To the best of our knowledge, this is the first time Senegal's airports have been analyzed using a satisficing DEA approach, in contrast to previous studies of this sector (Damacena et al., 2016; Wanke et al., 2016a). Thirdly, the present analysis includes an assessment of the impact of operational scope and different cost structure variables related to labor, capital, and cost-asset ratios on a given performance threshold.

Thus, the purpose of this study is to assess the determinants of efficiency within the context of Senegal's airports, based on cost structure variables commonly found in the literature. In order to achieve this objective, an efficiency analysis is developed using a two-stage approach: satisficing DEA model efficiency distributions are computed first, followed by SVM regression. The paper is structured as follows: after this introduction, the contextual setting is presented, including a description of Senegal's airports. The literature survey is then presented, followed by the methodology section, in which the two-stage satisficing DEA/SVM regression is further discussed. Section 5 presents the data, followed by a discussion of the results and the conclusion in Section 6.

2. Contextual setting

Senegal, located on the West African coast near the Gulf of Guinea, is a former French colony that became independent in 1960. Since then the country began its development. Airports are part of development infrastructure. The country has around 20 airports, but in this paper we focus on the country's 5 main airports: the capital, Dakar's, airport, followed by regional airports in main cities such as San Louis, Tambacounda, Ziguinchor, and Cap Skiring. All these airports have regular traffic leveraged on Senegal's development and population (15 million people in 2015). French is the common language and public administration follows the French tradition with an airport regulatory agency. The importance of airports in the country is due to the country geographical characteristics, with much of the northern part of Senegal's coast covered by dunes from Cap Vert to Saint-Louis, but with low rainfall as it is a desert area, and the southern part of Senegal composed of muddy estuaries with heavy rainfall. In the hinterland a sandy plain extends north to the floodplain of the Senegal River. Therefore, air travel is the most efficient mode of transportation. Table 1 presents some characteristics of Senegal's airports.

As can be seen in Table 1, the capital's airport is the most important in terms of all attributes, followed by Ziguinchor airport. The other airports do not have cargo operations and San Louis airport is the smallest in terms of traffic. Senegal's airports are a main asset of the country's infrastructure and an instrument of the country's development.

3. Literature review and research motivations

A recent review of airport efficiency papers, depicting the country of origin, the models applied and the variables used can be found in Damacena et al. (2016) and Wanke et al. (2016a). The usual sample size ranges from 11 to 67 airports while most studies relied on a single year or up to three or four-year data panels at the country level. Airport performance is usually analyzed in terms of efficiency or productivity. DEA models are used in productivity and efficiency studies (Gillen and Lall, 1997; Adler and Berechman, 2001; Barros and Dieke, 2007; Barros et al., 2011), while SFA – Stochastic Frontier Analysis – models are usually adopted for overall productivity and efficiency performance assessment (Barros and Sampaio, 2004; Barros, 2008, 2009; Diana, 2010). Although European and US airports are frequently analyzed, African ones are rarely assessed (Barros and Marques, 2010; Barros, 2014; Damacena et al., 2016; Wanke et al., 2016a).

A bibliometric analysis on the inputs and outputs used in the 27 different airport efficiency studies presented in Damacena et al. (2016) and Wanke et al. (2016a) reveals the most common ones. Specifically with respect to the inputs used, there were 95 nominations. Among them, the most frequent ones were (i) employees

Table 1
Characteristics of Senegal's airports in 2015.

Airports	Runway length (ft)	Passengers	Cargo	Aircraft Movements	Personnel Employed
Dakar	11 450	2 234 331	29 830	42 290	263
San Louis	6 372	3 202	0	811	11
Tambacounda	6 562	1 102	0	218	7
Ziguinchor	4 413	42 538	500	2 492	105
Cap Skiring	4 757	28 000	0	1 292	53

(number, cost, or payroll) with 20% of the total input nomination; (ii) runway (number or length), with 18.9%; (iii) terminal or airport area, with 16.8%; and operating costs or expenses, with 7.4%. On the other hand, these same 27 studies presented 85 output nominations. The most frequent ones were: (i) passenger throughput (27.1%); (ii) landing and take-offs or aircraft movements (25.9%); and (iii) cargo throughput (24.7%). Therefore, as further discussed in Section 5, this research uses the two most frequently nominated input variables and the three most frequently nominated output variables for describing airport efficiency as found in the literature review. Except for the apron area, which is an specific area for parking and maintenance, this input/output set is similar to that used in Curi et al. (2011).

In a broad sense, according to Yu (2010), airport efficiency studies seek to evaluate comparative operational efficiencies; illustrate how efficiency measures can be useful in monitoring airport operations; identify characteristics or context variables that may explain differences in airport operational efficiency; assess scale-impact on efficiency levels; and, measure production slack. Yu (2010) affirms, however, that context variables affecting airport efficiency should not be neglected as policy implementation may be unclear to airport authorities.

A special emphasis is placed in this research on understanding the impact of cost structure variables related to labor, capital, and cost-asset ratio on a given efficiency threshold. The idea is to measure whether higher efficiency levels in airport operations are directly related to certain aspects of the cost structure when compared to lower efficiency airports. The literature review indicates that the airports' underlying production technology is an important issue, with several different efficiency models attempting to capture its main features. This can be inferred not only from the fact that the size of the airport is relevant for achieving higher efficiency levels (technology may vary depending upon scale) but also due to the fact that the operational focus may also have an impact on efficiency levels, i.e., airports that specialize in passengers may exhibit efficiency levels that are different from those that also have cargo operations. Similarly, well-trained and remunerated employees may also reflect higher efficiency levels. This being the case, this research proposes to study the technology impacts of a given Senegal airport on a given efficiency threshold in terms of its cost structure. The idea is to understand how variables such as the capital-labor ratio and cost-asset ratio are reflected in the efficiency levels achieved by the airports in Senegal and their potential for deriving policy-making.

Lastly, based on the literature review, we verified that, thus far, no paper has simultaneously adopted satisficing DEA and SVM regression approaches; moreover, no important paper has undertaken an analysis of Senegal airports using this methodological context, which constitutes an additional novelty of this empirical study. It is also worth noting that there is a growing trend in the literature towards adopting two-stage analyses, whereby DEA scores are first computed and then undergo multivariate data analysis for correlation with a set of explanatory variables.

4. Satisficing DEA

DEA is a non-parametric model first introduced by Charnes et al. (1978). Based on linear programming (LP), it is used to address the problem of calculating relative efficiency for a group of DMUs by using a weighted measure of multiples inputs and outputs (Wanke, 2012a, 2012b; Kruger et al., 2002). Consider a set of n observations on the DMUs (Decision Making Units). Each observation, DMU_j ($j=1,\dots,n$) uses m inputs x_{ij} ($i=1,\dots,m$) to produce s outputs y_{rj} ($r=1,\dots,s$). DMU_o represents one of the n DMUs under evaluation, and x_{io} and y_{ro} are the ith input and rth output for DMU_o, respectively. Model (1) presents the envelopment modeling for the variable return-to-scale frontier types, where ε is a non-Archimedean element and s_i^- and s_r^+ account, respectively, for the input and output slack variables (Zhu, 2003; Bazargan and Vasigh, 2003).

$$\begin{aligned}
&\max \phi - \varepsilon\left(\sum_{i=1}^{m} s_i^- + \sum_{r=1}^{s} s_r^+\right)\\
&s.t.\\
&\sum_{j=1}^{n} \lambda_j x_{ij} + s_i^- = x_{io},\ \forall i\\
&\sum_{j=1}^{n} \lambda_j y_{rj} - s_r^+ = \phi y_{ro},\ \forall r\\
&\lambda_j \geq 0,\ \forall j\\
&\sum_{j=1}^{n} \lambda_j = 1
\end{aligned} \tag{1}$$

Any of the DMUs may or may not be on the frontier when the output-input ratio is measured (Barros and Peypoch, 2009; Wang et al., 2012; Wang and Feng, 2015). The distance from the actual allocation of a particular DMU to the frontier is believed to represent the inefficiency of the DMU, which may be caused by various factors that are specific to the DMU. If the efficiency of DMU i is 1, DMU i is a technically efficient DMU; if its efficiency is less than 1, it is technically inefficient.

Charnes and Cooper (1959) were the first to propose chance-constrained programming to measure efficiency under uncertainty and which analyses the cases of the possibility of violated constraints. Thore (1987), Banker (1993), and Land et al. (1993, 1994) made efforts to address data uncertainty in terms of stochastic variation in DEA. To accommodate stochastic variation, the constraint equations of the model (1) are modified and the mechanism of the chance-constrained formulation introduced by Land et al. (1993) is applied. Thus, the corresponding chance-constrained efficiency measure is calculated in line with the constant returns-to-scale assumption as:

$$\begin{aligned}
&\text{Max } \phi\\
&\text{subject to}\\
&Prob\left[\sum_{j=1}^{n} y_{rj}\lambda_j \geq \phi y_{ro}\right] \geq \alpha_r,\ \ r = 1,2,\dots,s,\\
&Prob\left[\sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{io}\right] \geq \alpha_i,\ \ i = 1,2,\dots,m,\\
&\lambda_j \geq 0,\ \ j = 1,2,\dots,n.
\end{aligned} \tag{2}$$

Here, "Prob" means probability and "~" identifies these inputs as random variables with known probability distributions (say, Normal, Beta, Gamma etc). The equality $\alpha_i = \alpha_r = 0.95$ is also assumed so that most DMUs (say 5%) will be set as best performers. We assume that the inputs and outputs are stochastically independent; the performance of one airport is independent of that of another airport. In order to extend the potential uses of the DEA models to cases where 100% efficiency can be replaced by performance aspiration levels, Cooper et al. (1996) incorporated Simon's (1957) satisficing concepts into the DEA models with chance constraints. In line with Udhayakumar et al. (2011) and with the support of the above literature, the probabilistic chance-constrained DEA model with "satisficing" concepts incorporated can be defined for model (2) as follows:

$$\begin{aligned} &\text{Max } P(\phi \geq \gamma) \\ &\text{subject to} \\ &Prob\left[\sum_{j=1}^{n} y_{rj}\lambda_j \geq \phi y_{ro}\right] \geq \alpha_r, \quad r = 1,2,\dots,s, \\ &Prob\left[\sum_{j=1}^{n} x_{ij}\lambda_j \leq x_{io}\right] \geq \alpha_i, \quad i = 1,2,\dots,m, \\ &\lambda_j \geq 0, \quad j = 1,2,\dots,n. \end{aligned} \tag{3}$$

Here, "Prob" and "~" are as defined above and one can interpret $\beta(=\gamma^{-1})$ as a performance threshold either imposed by an outside airport authority or adopted by a decision maker for some kind of analysis. It is to be noted that in model (3) the satisficing level is imposed only to the objective function and not at the constraint level, which means that the satisficing level at the constraint level is fixed at 100%. As a matter of fact, if $\beta = 1$, DMU_0 is called stochastically efficient if and only if $Prob(\phi \geq \gamma) = \alpha_o$. On the other hand, if $\beta < 1$, DMU_0 is called Satisficing-efficiency if and only if $Prob(\phi \geq \gamma) = \alpha_o$. Table 2 provides the pseudo code for the bootstrapped random generation of the inputs and the outputs, with external constraint check.

5. Data and efficiency assessment

5.1. The data

The data on five Senegal airports was obtained from the Agence Nationale de L'Aviation Civile du Sénégal which regulates Senegal's airports and covers the 1996–2015 period. The choice of inputs and outputs is perhaps the most important task in employing DEA to measure the relative efficiency of DMUs. The inputs and outputs considered were chosen not only because they were commonly found in the literature review but also in accordance with the availability of data regarding physical productive resources. The two input variables considered included human and physical resources related to airport operations (1. Number of Employees; and 2. Runway length in feet). On the other hand, the three output variables considered included different measures of airport production (1. Cargo movement in tonnes per year; 2. Number of aircraft movements per year; and 3. Number of passengers per year). Their descriptive statistics are presented in Table 3. In addition to these inputs and outputs, it should be noted that cost structure variables, such as 1. Cost of capital (measured as the amortization-asset ratio); 2. Cost of labor (measured as the wages-number of employees ratio); 3. Cost-asset ratio, and 4. Whether the airport has cargo operations as well as passenger traffic (1 = yes; 0 = no). The idea is to control the computed efficiency scores for the cost structure of airports, whether capital or labor intensive, while focusing on their operational diversity or scope. Monetary variables used in this research to compute the cost structure variables were expressed in current thousand USD, adjusted for this country's annual inflation. Their descriptive statistics are also presented in Table 3.

Table 2
Pseudo code.

1. Adjust the corresponding probability distribution for the inputs and outputs of the data set
2. Run N – a sufficiently large number – times. In this research, N = 100
   a. Generate artificial data by random numbers in accordance with each corresponding probability distribution
   b. Obtain the efficiency score by solving the probabilistic chance constrained DEA with the chosen performance thresholds for every realization
   c. Record efficiency scores
3. Compute relevant Statistics
4. Finish.

Before proceeding, it is worth testing the underlying assumption of a possible distributional fit for the original input and output set against an eventual change in the productivity of Senegal airports over the course of these twenty years. Therefore, a robustness analysis was performed in terms of Malmquist Index (MI), allowing the temporal decomposition of the productive change (MI) in its two major components: efficiency change (or catch-up effect) and technical change (or frontier shift effect). Still, it is noteworthy that information on the homogeneity assumption of DMUs also provides guidance on this robustness analysis. Readers should recall that the cargo movement for some airports is equal to zero. These zero values were substituted by 0.01—according to the feature offered by DEA softwares (Barr, 2004; Hwang et al., 2016) — in order to proceed with the analyses. Results suggest that the methodological bias introduced by this procedure seems to be minimal, as they still hold when 0.001 is used instead of 0.01 or even 0.

As suggested in Fig. 1, Senegal airports showed a stagnant pattern over the course of the time, with scores for the productive change, efficiency change, and technical change strongly concentrated around one. As a matter of fact, only the year 2000 was marked by a localized spike in the catch up and frontier shift effects, mostly concentrated in Dakar airport. This year coincided with the privatization of Air Senegal: fifty-one per cent of it was bought by Royal Air Maroc in January 2000, producing a self-contained boom at that time (OECD report on Senegal economy, accessed at: https://www.oecd.org/countries/senegal/1826266.pdf in September 12th, 2016). The overall implication of these results is that, since there is no acknowledgeable systematic trend in productivity and efficiency in Senegal airports over the course of the years, it is possible to consider their input-output vector as random variates. Additionally, a smoothed bootstrapped MI, following the steps presented in Fuentes and Lillo-Bañuls (2014) was performed to remove the inherent bias of a small sample of five airports per each year and

Table 3
Descriptive statistics for inputs, outputs, and contextual variables.

Variables		Min	Max	Mean	SD	CV
Inputs	Personnel	5	270	77.02	83.87	1.09
	Runway length (ft)	4 413	11450	6704.66	2519.49	0.37
Outputs	Passengers	837	2 234 331	349174.5	689 942.3	1.98
	Cargo ([a])	0	29830	4695.93	9420.76	2.01
	Aircraft	133	42290	7317.72	12 935.33	1.77
Contextual	Cost Asset Ratio	0.02	6.09	0.98	1.65	1.68
	Cost of Capital	0.02	23.35	0.89	2.55	2.85
	Cost of Labor	74.93	700.73	241.85	139.29	0.58
	Cargo Operation	Yes:	40%	No:	60%	

[a] Minimal values equal to zero indicate cases where the airport does not operate cargo traffic.

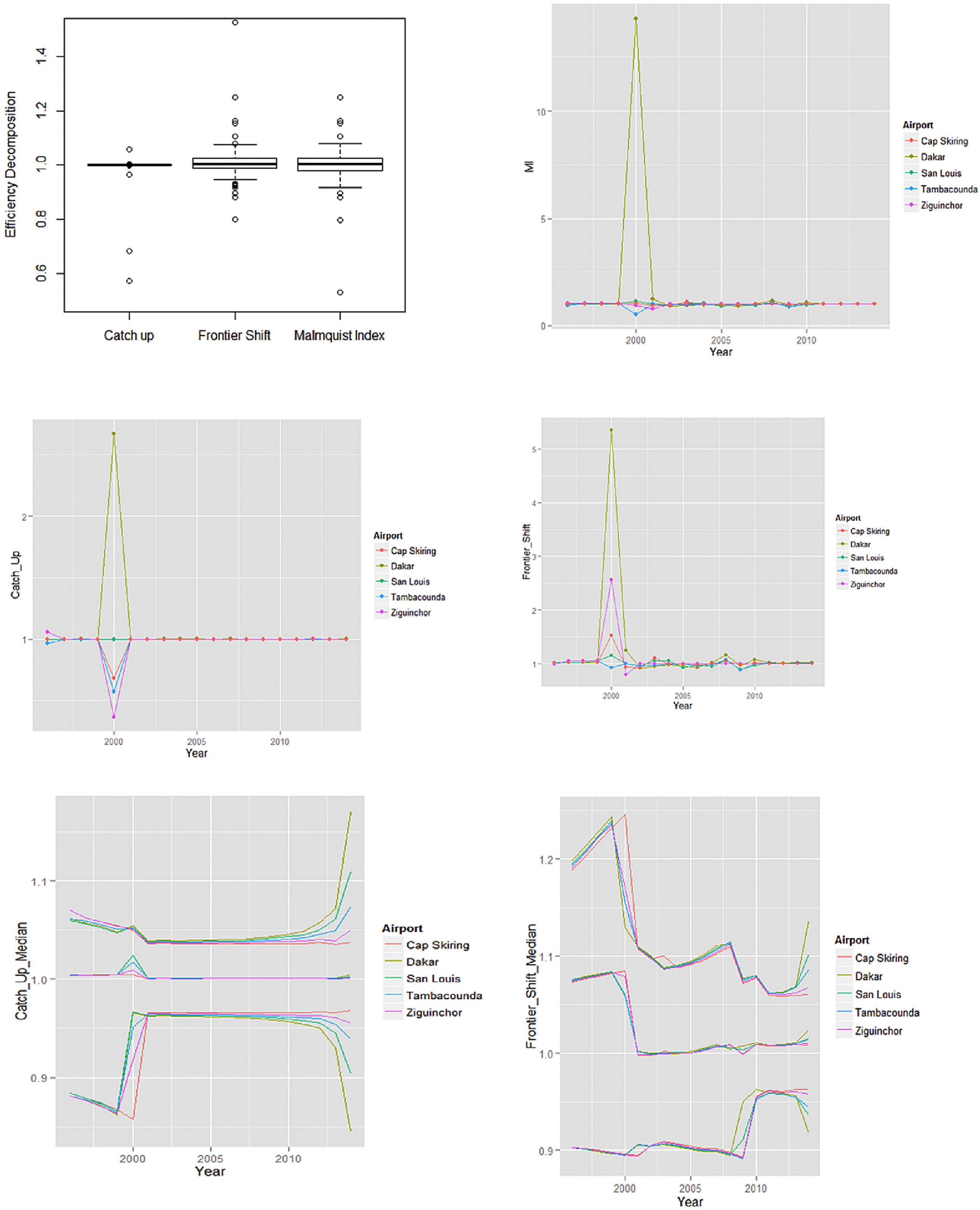

Fig. 1. Robustness results for the MI of Senegal Airports (1996–2015) – top and middle. Results for the smoothed bootstrap MI of Senegal Airports (1996–2015) – bottom.

corroborate these findings. 1000 replications were performed for each year. Fig. 1 (bottom) presents the 95% confidence intervals for the bootstrapped efficiency scores. It also indicates that it is not possible to conclude in favor of a systematic increase in productive change, efficiency change, and technical change over the course of the years, since both lower and upper confidence limits are either below one or above one, respectively. Readers should also pay attention to the difference in scale between Fig. 1 middle and bottom, as result of bias removal. Bootstrapped MI indexes were, therefore, considerably deflated.

Therefore, the subsequent step to the collection of the inputs and outputs was their adjustment to best continuous distribution for each DMU. The goodness of fit procedure used in this research followed two steps: first, an indicative approach to the best fit by the Cullen & Frey graph, showed in Fig. 2 for each input and output. Second, a Maximum Likelihood Estimation (MLE) to derive the parameters of the chosen distributions in the first step (Delignette-Muller and Dutang, 2014). Results are not available for those inputs and outputs that are constant over time, such as runway length and cargo movement at some airports.

The Cullen & Frey graph is a skewness-kurtosis plot such as the one proposed by Cullen and Frey (1999). In this plot, common distribution values are displayed in order to help the choice of distributions fit the data. For some distributions (i.e. normal, uniform, logistic, exponential), a single point on the plot represents the distribution, because there is only one possible value for the skewness and the kurtosis. For other distributions, areas of possible values are represented, consisting of lines (i.e. gamma and lognormal), or larger areas (i.e. beta). Delignette-Muller and Dutang (2014) warn that skewness and kurtosis are not considered to be robust, due to their high variance, and suggest a nonparametric bootstrap procedure in order to take into account the uncertainty of the estimated values of the data's degree of kurtosis and skewness (Efron and Tibshirani, 1994). A resample size of 100 was adopted here to generate the yellow dots presented in Fig. 2.

Once selected, one or more parametric distributions may be fitted to the data set, one at a time. The distribution parameters were estimated by maximizing the likelihood function, considering the observations of each criterion and the density function of the parametric distribution. Numerical results returned the parameter estimates, the estimated standard errors (computed from the estimate of the Hessian matrix at the maximum likelihood solution), the log likelihood, the Akaike and Bayesian information criteria (the so-called AIC and BIC), and the correlation matrix between parameter estimates. Table 4 summarizes the best fitting probability distributions and their estimated parameters for each input and output at each airport (Delignette-Muller and Dutang, 2014). Once again, results are not available for those inputs and outputs that are constant over time. Results suggest that beta distribution exhibited the best fit for almost all variables, with the exception of cargo movement for two airports that adhered to a normal distribution. These distributions and the respective parameters were used in the Satisficing DEA model presented in Section 4. Inputs and outputs were rescaled in a unity-based normalization for each airport, bringing all values into the range (0, 1).

The temporal perspective of this rescaled input-output set per each airport does not corroborate, in most cases, the presence of a systematic increasing or decreasing trend but rather suggests a random movement around an average when the long-term picture is analyzed. Besides, the expected beneficial impact of Air Senegal

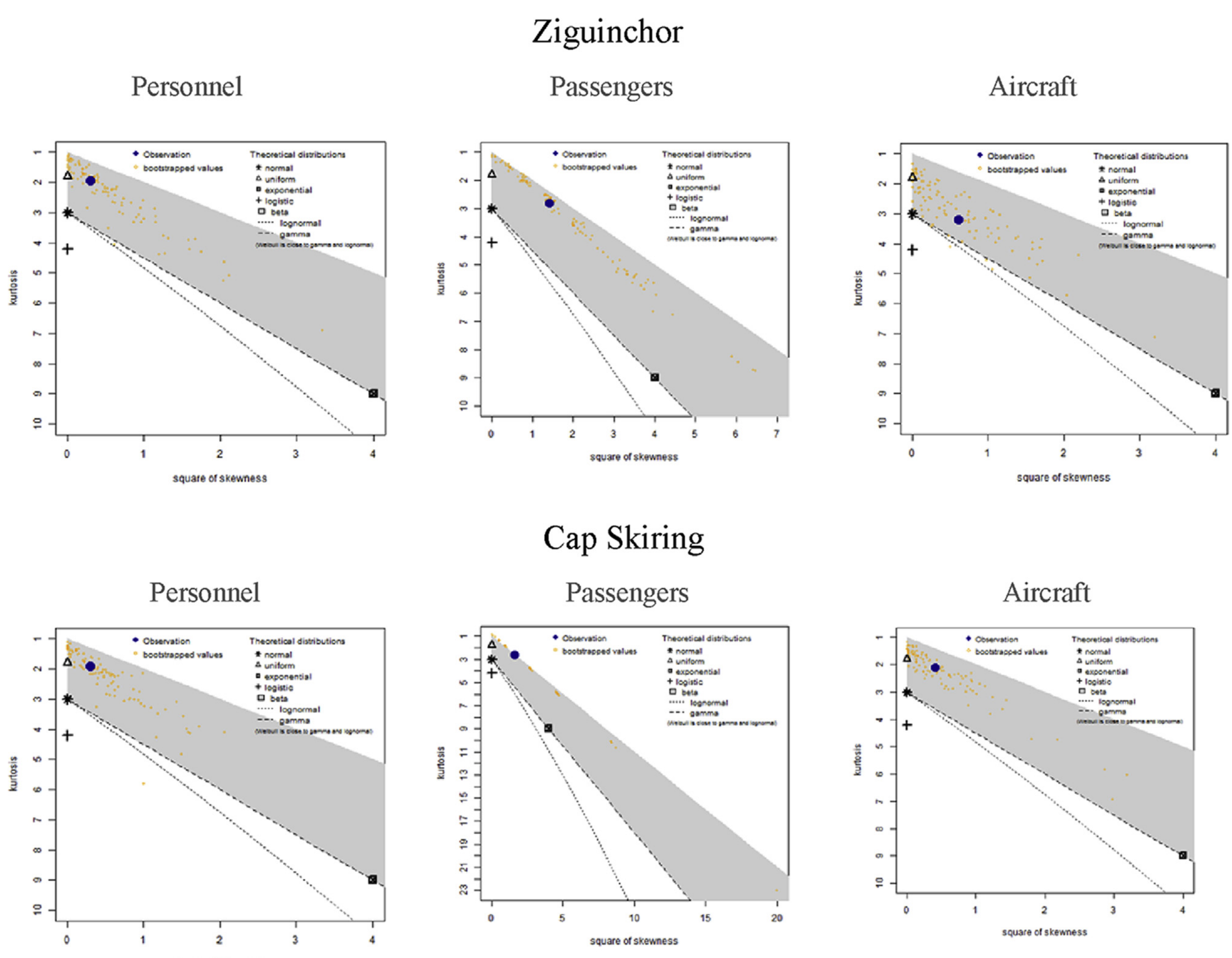

Fig. 2. Cullen & Frey graphs for the inputs and outputs of Senegal's airports.

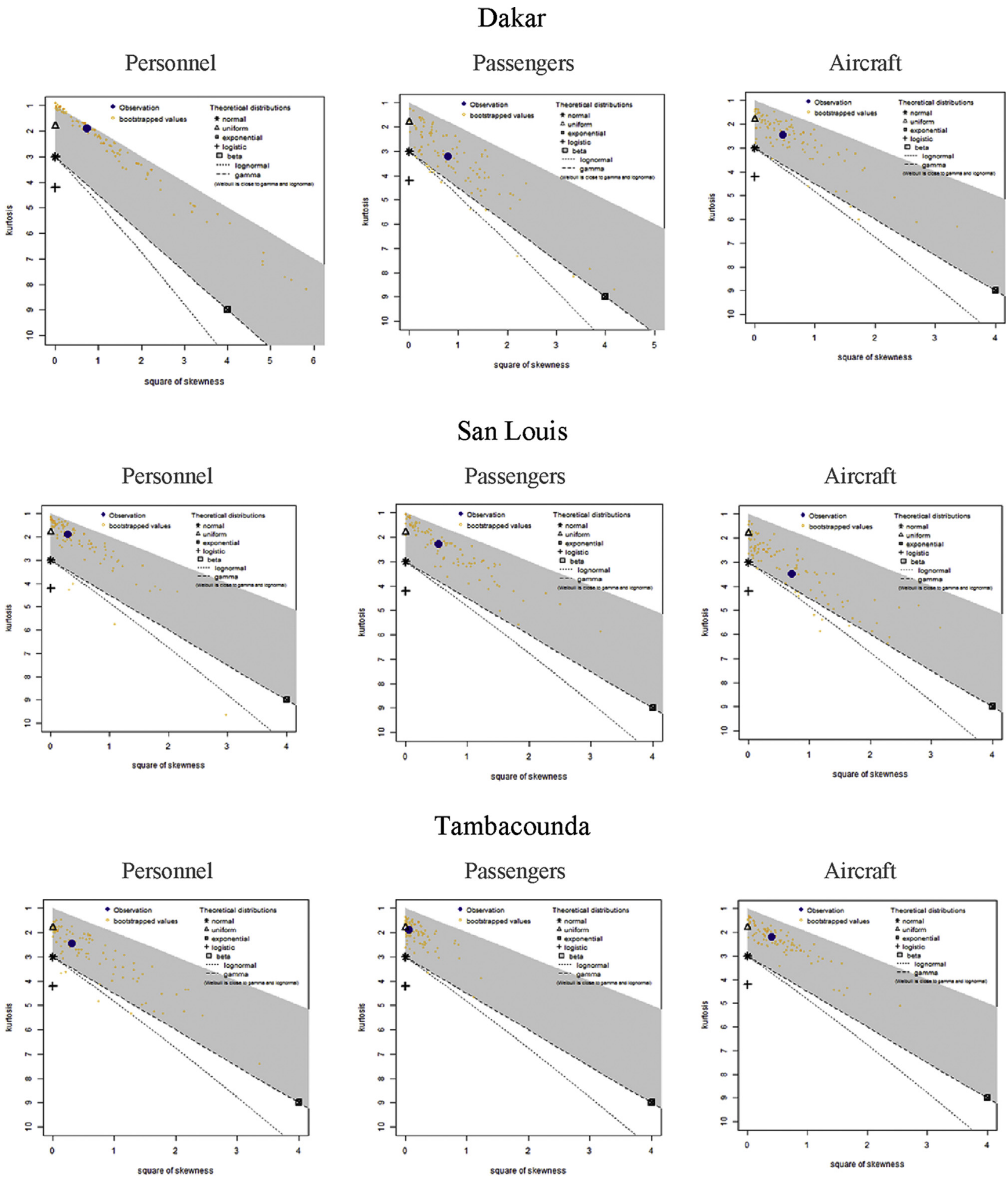

Fig. 2. (*continued*).

privatization on Senegal airports, often reflected as a spike in the year 2000, is dissipated in the long term since random fluctuations appear to be the dominant pattern from that time on. Nevertheless, this being the case, it is possible to affirm, that the increase in personnel employed due to route restructuring occurred in a more than proportional fashion when compared to the increase in passenger traffic and aircraft movements (cf. Fig. 3).

5.2. Discriminating between high/low efficiency groups within each performance threshold using Support Vector Machine regression

Support Vector Machines are predictive techniques that have been receiving increased attention from different research communities, due to their successful application in several domains, in addition to their strong theoretical background (Torgo, 2011).

Table 4
Outcomes for the best distributional fit ([a]).

Airports	Number of employees	Runway length	Cargo movement	Aircraft movement	Passenger movement
Dakar	B(0.389, 0.287)	NA	N(0.488, 0.248)	B(0.379, 0.427)	B(0.321, 0.369)
San Louis	B(0.264, 0.333)	NA	NA	B(0.389, 0.561)	B(0.409, 0.369)
Tambacounda	B(0.514, 0.410)	NA	NA	B(0.407, 0.512)	B(0.422, 0.463)
Ziguinchor	B(0.422, 0.349)	NA	N(0.280,0.356)	B(0.364, 0.514)	B(0.258, 0.386)
Cap Skiring	B(0.341, 0.329)	NA	NA	B(0.358, 0.475)	B(0.154, 0.298)

[a] NA – not available - indicates that the variable either did not exhibit any variation over time at that specific airport (runway length) or was zero (as in cargo movement, thus implying that the respective airport does not operate cargo).

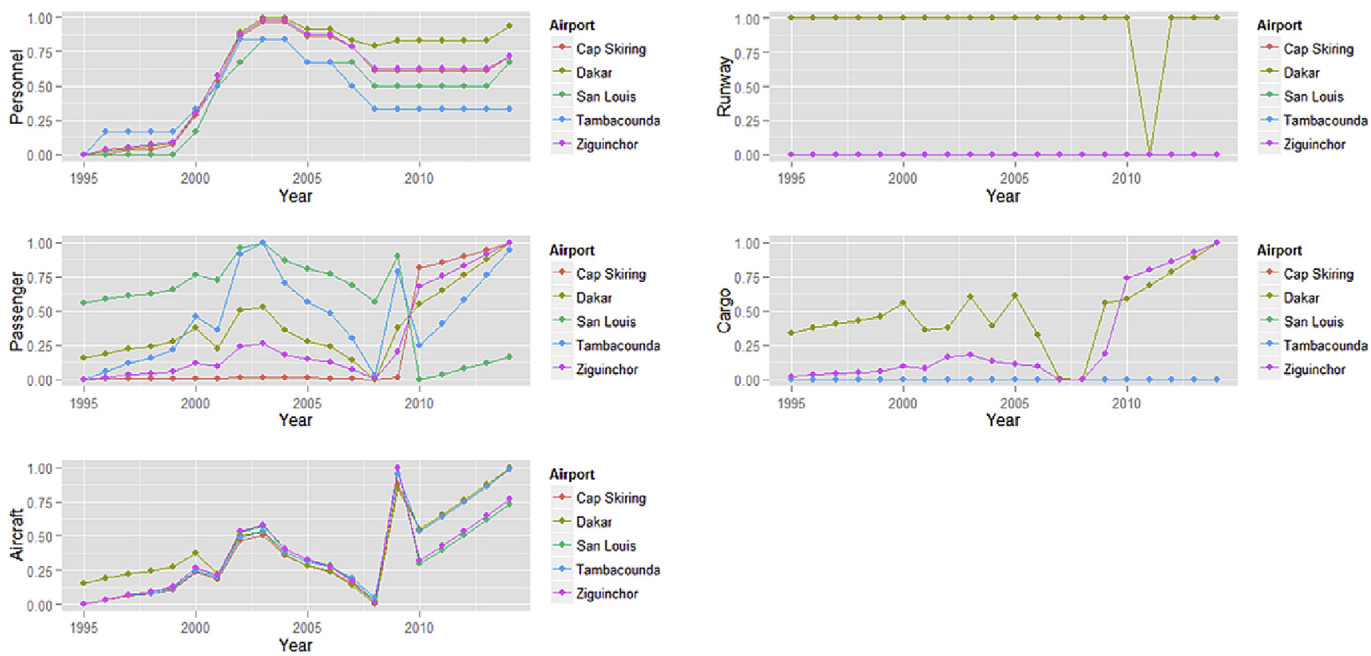

Fig. 3. Rescaled inputs and outputs in Senegal airports.

Vapnik (1995, 1998) and Shawe-Taylor and Cristianini (2000) are two fundamental references for SVMs. The basic idea behind SVMs is that they map the original data into a new, high dimensional space, where it is possible to apply linear models to obtain a separating hyper-plane (James et al., 2013; Ledolter, 2013). The mapping of the original data into this new space is carried out with the help of so-called kernel functions (Torgo, 2011). Originally intended for the binary classification setting in which there are two classes, SVMs are closely related to statistical methods such as logistic regression (James et al., 2013). In this research, in order to discriminate between high/low efficiency groups within each performance threshold, SVM regression analyses were performed considering the contextual variables as the predictor variables.

The aim of the SVM regression is to find a nonlinear generalist function that describes the suggested model with minimum error. Thus, the SVM proposes a hyper-plane in which data in the training base are the closest together possible (Beltrami, 2009; Drucker et al., 1997). In addition, this version of Support Vector Machines allows us more freedom to choose penalization parameters and the degree of flattening of the function, thus obtaining better fit.

The basic problem of regression is to find a function that fits a specific set of data. A function ((??)) must be found that fits a specific vector ?? at less than a specified error ?? This involves using a measure of the degree of loss known as a Loss Function, described as |?? − (??)| ≥ ??

In addition to considering the loss function, the SVR also attempts to minimize the reciprocal of the margin and allows the inclusion of errors. According to Soman et al. (2011), the primal formulation of the SVM regression may be described as:

$$
\begin{aligned}
&\text{Min } W = \frac{1}{2} w^T w + C\xi^T 1 + C\xi^{*T} 1\\
&\text{Subject to :}\\
&\phi^T(x_i) w + \xi_i^* \geq -\varepsilon 1 + y_i\\
&-\phi^T(x_i) w + \xi_i \geq -\varepsilon 1 - y_i\\
&\xi_i \geq 0\\
&\xi_i^* \geq 0
\end{aligned}
\tag{4}
$$

The primal problem of the SVM regression may also be written in its dual form, given by the Wolfe (1961):

$$
\begin{aligned}
&\text{Max } Z = -\frac{1}{2}\left(\lambda^{T*} - \lambda\right) K \left(\lambda^* - \lambda\right) - \varepsilon\left(\lambda^{T*} 1 - \lambda^T 1\right) + \left(\lambda^{T*} - \lambda\right) y\\
&\text{Subject :}\\
&\lambda^T 1 - \lambda^{T*} 1 = 0\\
&0 \leq \lambda \leq C1^T\\
&0 \leq \lambda^* \leq C1^T
\end{aligned}
\tag{5}
$$

where K represents the kernal function; C is the cost parameter attributed to the classification or prediction error, i.e., it is the penalization; ?? represents the vector of the element one wishes to predict, the variable that will be explained by the models, containing all the observations in a base that will be used to train the machine and ?? is the prediction error – also called bias term. The matrix ?? contains all the variables and observations considered to explain ??; ε is the parameter of the model's margin of error which allows some leeway so that the resulting function is not overly-fitted to the training data, as this, according to Soman et al. (2011) and Morettin (2014), would probably cause poor prediction performance.

In SVM regression, the strategy for dealing with nonlinearity in data is to create new dimensions using a mapping process, which is described as follows:

$$\begin{array}{l} x \rightarrow \phi\ x \\ \mathbb{R}^p \rightarrow \mathbb{R}^q \text{ given that } q \gg p. \end{array} \quad (6)$$

There are various mappings that can lead to different space characteristics and the challenge that arises is precisely to identify which is the best one for a specific classification problem in order to minimize the generalization error.

6. Results and discussion

Initially, traditional DEA constant returns-to-scale (CRS also known as CCR model) estimates revealed the existence of 67 efficient airport observations between 1996 and 2015. These respective efficiency estimates for the whole sample, using the model presented in eq. (1) – dropping off the constraint of summing one - are given in Fig. 4. The mean overall efficiency scores in the traditional CRS DEA method is 0.892, whereas the traditional DEA varying returns-to-scale (VRS) model, also known as BCC model, exhibited a mean value of 0.947. This result suggests, as expected, that the discriminatory power of the traditional CRS model is higher than that observed in the VRS model, because their scores are lower and not so inflated towards one.

On the other hand, the chance-constrained DEA model based on bootstrapped inputs and outputs at the DMU level showed better discriminatory power, with median overall scores around 0.65. The mean efficiency ranking for each one of the five Senegal airports is as follows: Dakar (0.702); Cap Skiring (0.633); Ziguinchor (0.540); Tambacounda (0.495) and San Louis (0.484). Fig. 5 illustrates the behavior of a partial efficiency frontier for the Senegal airports, considering two rescaled outputs – passenger and aircraft – and one rescaled input – personnel. Observations concentrated in the top right part of the 3D plot suggest observations with higher efficiency levels, while those located at the bottom parts of the graph may denote lower efficiency levels. Readers should note that the distribution of the observations of the different airports is not homogeneous within these quadrants, thus affecting the final efficiency score computed for each airport.

The fact that the largest airport, Dakar, is the most efficient in Senegal and the very nature of the CRS model presented in

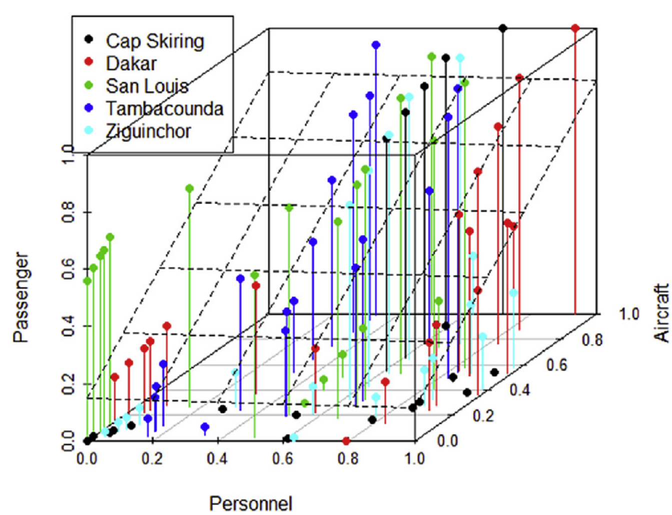

Fig. 5. Partial frontier 3D plot for the rescaled inputs and outputs in Senegal airports.

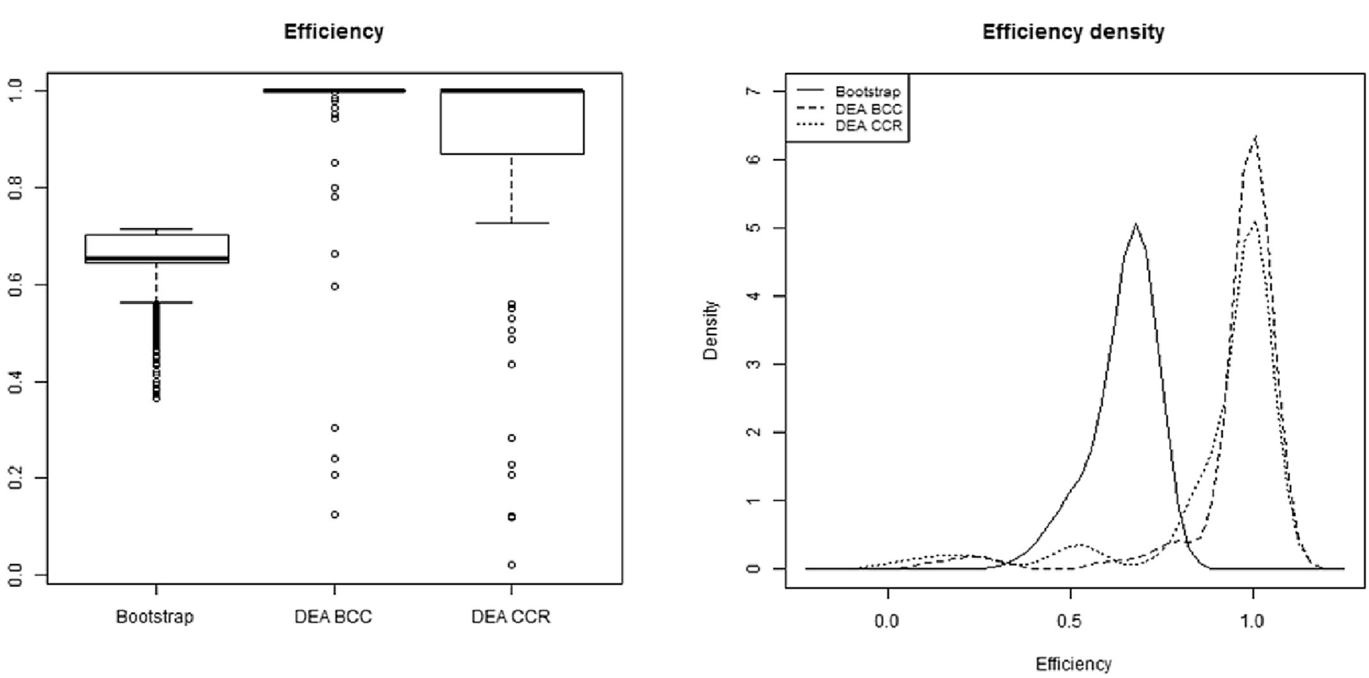

Fig. 4. Efficiency estimates distribution.

equations (2) and (3) indicates that efficiency is affected by scale aspects, besides managerial or technical efficiency. In this research, we assume that these managerial aspects are related to the quality of the human resources and their training, which are reflected in the cost of labor. As regards size, cost of capital and the cost asset-ratio proxies capital-intensive, larger, airports. Readers should recall that the CRS assumption leads to the computation of what is known as overall technical efficiency, which can be decomposed in pure technical efficiency (managerial) and scale efficiency (size) effects. The aim of SVM regressions is to analyze these countervailing effects embedded in the CRS assumption, within the ambit of Senegal airports, at a given performance threshold. These issues are further explored next.

Results for the satisficing DEA model, with bootstrapped inputs and outputs and simultaneously observing equations (2) and (3), are depicted in Fig. 6 for a performance threshold of 70%. That is, the probability of airport efficiency being higher or lower than 70% is considered as the cut-off point to assess the impact of the contextual variables. As a matter of fact, different threshold values were tested to assure the best discriminatory power between high/low efficiency groups, which was indeed obtained for the threshold of 70%. Readers should note that the cost of labor is more important in the higher efficiency group (threshold $\geq$ 70%) than in the lower efficiency group (threshold $<$ 70%). These results suggest that higher efficiency levels at Senegal's airports may be driven by human resource skills, which are, to some extent, reflected in the salaries paid and, therefore, in the cost of labor. These results are somewhat in line to the increase in labor force verified in Senegal airports after privatization of Senegal Air, as discussed in Section 5. On the other hand, the performance of lower efficiency airports is driven by cargo operations, the cost of capital, and cost asset-ratio, suggesting that economies of scale and scope may play an important role when efficiencies are low. Airport efficiency in Senegal, therefore, may be a result of human resource quality, operational

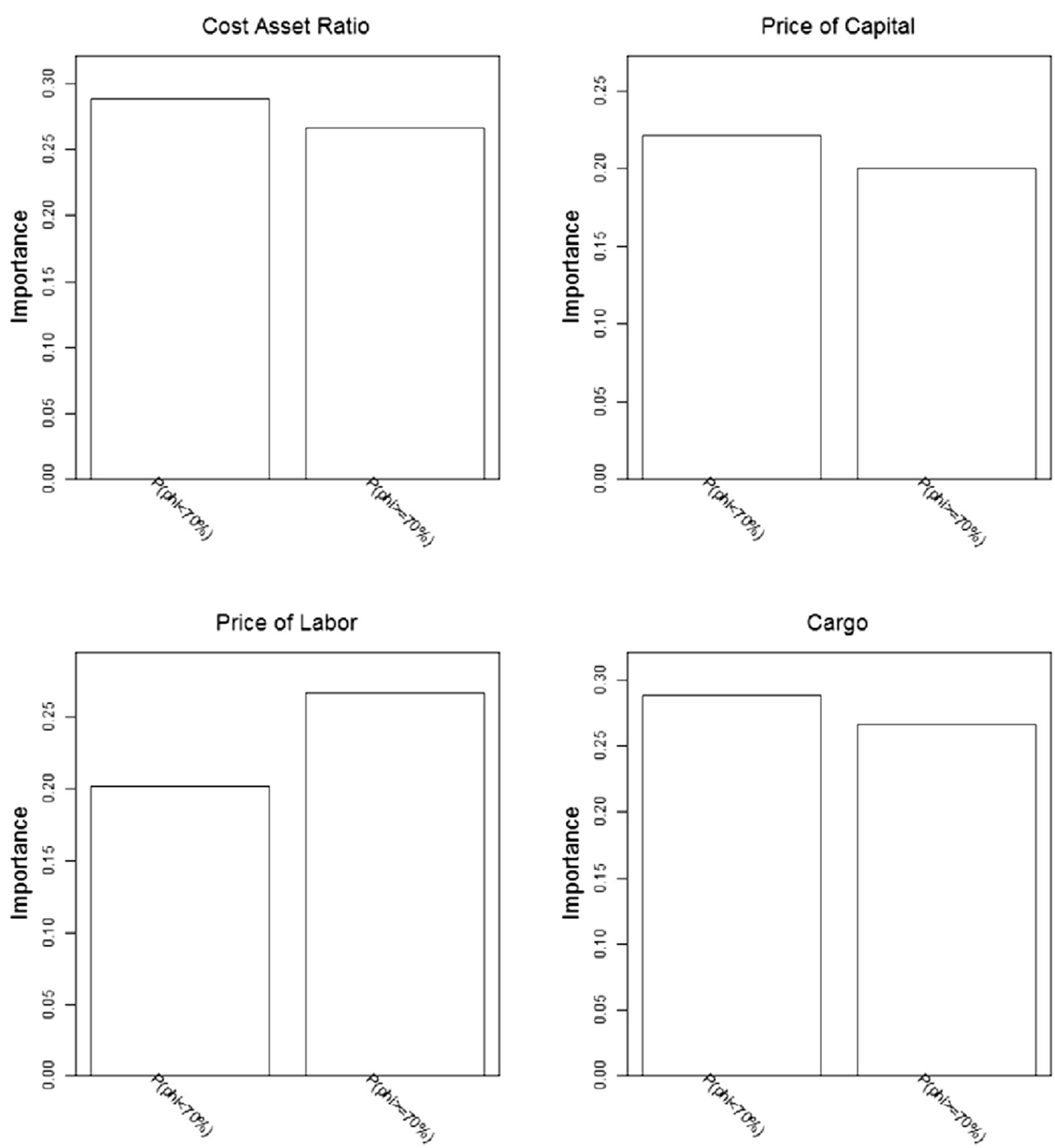

Fig. 6. Relative importance of the contextual variables given different performance thresholds or satisficing level using SVM regression.

scope and capital intensity, rather than simply a side effect driven by size.

The sensitivity analysis results of the SVM regression are presented in Fig. 6. They indicate the discrimination between the high/low efficiency groups formed considering the 70% performance threshold or satisficing level. The conditional mean efficiency was chosen as the cut-off value to split these high/low groups at the 70% threshold. More precisely, Fig. 7 presents the sensitivity analysis of the contextual variable estimates based on several bootstrapped SVM regressions. Their performance estimates are given in Table 5, indicating a reasonable adjustment (mean AUC – Area Under the Curve – of 0.742). These Variable Effect Curves (VECs) were computed as prescribed in Cortez and Embrechts (2013). Results suggest that the probability of belonging to the high efficiency group at the performance threshold of 70% tends to decrease at airports with capital-intensive structures and diversified operations. The implication of these findings for Senegal's airports is that they should focus on operational practices and more skilled human

Table 5
Bootstrapped performance estimates for the SVM regression.

	Mean	SE
Accuracy	0.519	0.229
Specificity	0.563	0.307
Sensitivity	0.425	0.329
AUC	0.742	0.200
R library: e1071 Function: svm	Parametrization: scale = TRUE, type = NULL, kernel = "linear", degree = 3, gamma = 0.1666667, coef0 = 0, cost = 1, nu = 0.5, class.weights = NULL, cache size = 40, tolerance = 0.001, epsilon = 0.1, shrinking = TRUE, cross = 0, probability = TRUE, fitted = TRUE, seed = 1 L, na.action = na.omit	

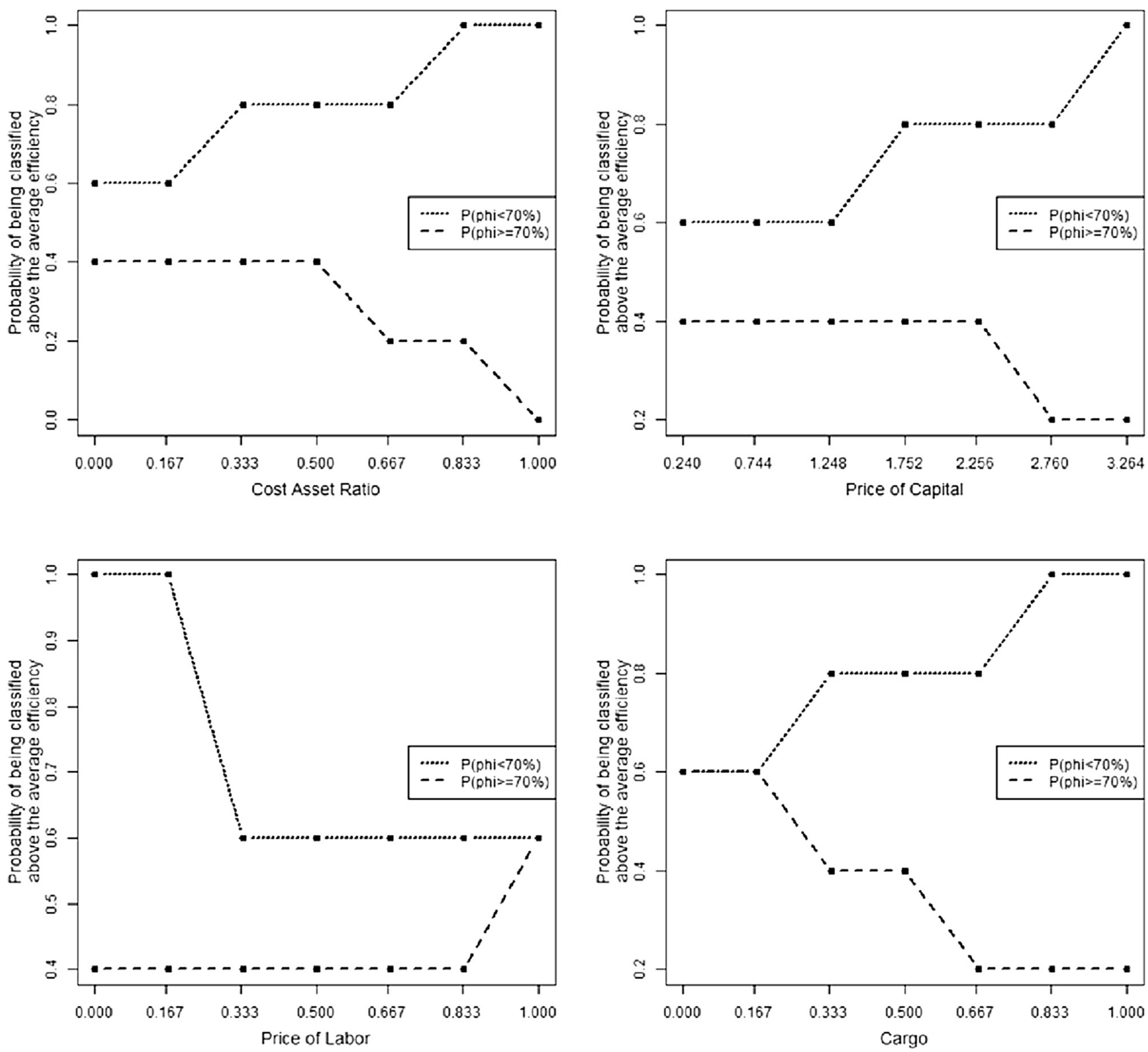

Fig. 7. Sensitivity analysis results for the SVM Regression.

resources beyond a certain efficiency level. This being the case, capital-intensive structures and diversified airport operations should be seen as key features when implementing new airport operations.

7. Conclusion

This paper presents an analysis of the efficiency of Senegal's airports using a satisficing DEA model and SVM regression. This approach has proven to be useful for handling small samples, since the discriminatory power is good. Satisficing DEA enables the efficiency of an airport to be assessed in terms of given performance thresholds, thus making it possible to identify how the impact of different contextual variables vary from lower to higher efficiency levels. SVM regression allows the discrimination of high/low efficiency airports at each given threshold based on conditional probabilities. Broadly speaking, based on the results, the most important contextual variables for Senegal's airport efficiency are those related to the capital-labor ratio and the diversity of operations, such as having cargo operations as well as passenger operations. This being the case, a greater emphasis should be given on human resource skills and training as well as on improving operational practices.

Scale and managerial style explain efficiency at Senegal's airports differently depending upon the efficiency threshold. The policy implication of this research is that Senegal's airports have to increase their efficiency in order to lower costs and increase quality of service. This ranking shows that complacency is the management practice at Senegal's airports and this paper is the first to present a comparison between the country's main airports. This result also shows that efficiency analysis should be adopted on a yearly basis, thus resulting in lower costs and greater quality. This procedure will enable airports to provide the population with a better service. With airports located in different places, comparisons have to be based on managerial practice, as reflected in balance sheets and income statements, similarly to the procedure used in this research. Given the importance of airports for the development of Senegal, further research is necessary to confirm these results, especially those related to other aspects of the cost structure. Other regions around the globe should also be the object of future studies. Limitations of this study are mainly related to the set of inputs and outputs used and to the data availability, which are intrinsic of secondary data.

References

Adler, N., Berechman, J., 2001. Measuring airport quality from the airlines' viewpoint: an application of data envelopment analysis. Transp. Policy 8 (3), 171–181.
Agpak, K., Gökçen, H., 2007. A chance-constrained approach to stochastic line balancing problem. Eur. J. Oper. Res. 180 (3), 1098–1115.
Banker, R.D., 1993. Maximum likelihood, consistency and data envelopment analysis: statistical foundations. Manag. Sci. 39 (10), 1265–1273.
Barr, R., 2004. DEA software tools and technology—a state-of-the-art survey. In: Cooper, W.W., Seiford, L.M., Zhu, J. (Eds.), Handbook on Data Envelopment Analysis. Kluwer Academic, Boston, MA, pp. 539–566.
Barros, C.P., 2008. Technical change and productivity growth in airports: a case study. Transp. Res. A 42 (5), 818–832.
Barros, C.P., 2009. The measurement of efficiency of UK airports using a stochastic latent front. Transp. Rev. 29 (4), 479–498.
Barros, C.P., 2014. Airports and tourism in Mozambique. Tour. Manag. 41, 76–82.
Barros, C.P., Dieke, P.U.C., 2007. Performance evaluation of Italian airports: a data envelopment analysis. J. Air Transp. Manag. 13 (4), 184–191.
Barros, C.P., Marques, R.S., 2010. Performance of Mozambiquean airports. Regul. Ownersh, Manag. Effic. 18 (1), 29–37.
Barros, C.P., Peypoch, N., 2009. An evaluation of European airlines' operational performance. Int. J. Prod. Econ. 122 (2), 525–533.
Barros, C.P., Sampaio, A., 2004. Technical and allocative efficiency of airports. Int. J. Transp. Econ. 31 (3), 355–377.
Barros, C.P., Wanke, P., 2015. An analysis of African airlines efficiency with two-stage TOPSIS and neural networks. J. Air Transp. Manag. 44–45, 90–102.
Barros, C.P., Managi, S., Yoshida, Y., 2011. Heterogeneity in the technical efficiency in Japanese airports. Singap. Econ. Rev. 56 (4), 523–534.
Bazargan, M., Vasigh, B., 2003. Size versus efficiency: a case study of US commercial airports. J. Air Transp. Manag. 9 (3), 187–193.
Beltrami, M., 2009. Precificação de opções sobre ações por modelos de Support Vector Regression. Dissertation. Universidade Federal do Paraná.
Bhattacharya, U.K., 2009. A chance constraints goal programming model for the advertising planning problem. Eur. J. Oper. Res. 192 (2), 382–395.
Charnes, A., Cooper, W.W., 1959. Chanced-constrained programming. Manag. Sci. 6 (1), 73–79.
Charnes, A., Cooper, W.W., 1963. Deterministic equivalents for optimizing and satisficing under chance constraints. Oper. Res. 11 (1), 18–39.
Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. Eur. J. Oper. Res. 2 (6), 429–444.
Chen, T., 2002. A comparison of chance-constrained DEA and stochastic frontier analysis: bank efficiency in Taiwan. J. Oper. Res. Soc. 53 (5), 492–500.
Cooper, W.W., Huang, Z., Li, S., 1996. Satisfying DEA models under chance constraints. Ann. Oper. Res. 66 (4), 279–295.
Cortez, P., Embrechts, M.J., 2013. Using sensitivity analysis and visualization techniques to open black box data mining models. Inf. Sci. 225, 1–17.
Cullen, A.C., Frey, H.C., 1999. Probabilistic Techniques in Exposure Assessment: a Handbook for Dealing with Variability and Uncertainty in Models and Inputs. Plenum Press, New York.
Curi, C., Gitto, G., Mancuso, P., 2011. New evidence on the efficiency of Italian airports: a bootstrapped DEA analysis. Socio-Economic Plan. Sci. 45 (2), 84–93.
Damacena, E.F., Wanke, P.F., Correa, H.L., 2016. Infrastructure expansion in Brazilian airports: slack analysis using a distance friction minimization approach. Decis. 43 (2), 181–198.
Delignette-Muller, M.L., Dutang, C., 2014. Fitdistrplus: an R package for fitting distributions. J. Stat. Softw. 64, 1–34.
Diana, T., 2010. Can we explain airport performance? A case study of selected New York airports using a stochastic frontier model. J. Air Transp. Manag. 16 (6), 310–314.
Drucker, H., Burges, C.J.C., Kauffman, L., Smola, A., Vapnik, V., 1997. Support vector regression machines. In: Mozer, M.C., Joradn, J.I., Petsche, T. (Eds.), Neural Inform. Processing Syst. 9. MIT Press, Cambridge, MA, pp. 155–161.
Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap. CRC Press, Boca Raton.
Fuentes, R., Lillo-Bañuls, A., 2014. Smoothed Bootstrap Malmquist Index Based on DEA Model to Compute Productivity of Tax Offices, Expert Systems with Applications. http://dx.doi.org/10.1016/j.eswa.2014.11.002.
Gillen, D., Lall, A., 1997. Non-parametric measures of efficiency of US airports. Int. J. Transp. Econ. 28 (3), 283–306.
Hwang, S.-N., Lee, H.-S., Zhu, J., 2016. Handbook of Operations Analytics Using Data Envelopment Analysis. Springer, ISBN 978-1-4899-7705-2.
James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. An Introduction to Statistical Learning. Springer, New York.
Kall, P., 1976. Stochastic Linear Programming. Springer Verlag, Berlin.
Kruger, H.A., Steyn, P.J., Kearney, W., 2002. Determinants of internal audit efficiency. S. Afr. J. Bus. Manag. 33 (3), 53–62.
Land, K.C., Lovell, C.A.K., Thore, S., 1993. Chance constrained data envelopment analysis. Manag. Decis. Econ. 14 (6), 541–554.
Land, K.C., Lovell, C.A.K., Thore, S., 1994. Productivity efficiency under capitalism and state socialism: an empirical inquiry using chance-constrained data envelopment analysis. Technol. Forecast. Soc. Chang. 46 (2), 139–152.
Ledolter, J., 2013. Data Mining and Business Analytics. Wiley, New Jersey.
Li, S., Jahanshahloo, G.R., Khodabakhshi, M., 2007. A super-efficiency model for ranking efficient units in data envelopment analysis. Appl. Math. Comput. 184 (2), 638–648.
Merkert, R., Pearson, 2015. A non-parametric efficiency measure incorporating perceived airline service levels and profitability. J. Transp. Econ. Policy 49 (2), 261–275.
Morettin, P.A., 2014. Ondas e ondaletas: Da análise de Fourier à análise de ondaletas de séries temporais, 2th ed. Edusp, São Paulo.
Shawe-Taylor, J., Cristianini, N., 2000. An Introduction to Support Vector Machines. University Press, Cambridge.
Simon, H.A., 1957. Models of Man, Social and Rational. John Wiley & Sons, New York.
Soman, K.P., Loganathan, R., Ajay, V., 2011. Machine Learning with SVM and Other Kernel Methods. PHI Learning Private Limited, New Delhi.
Sueyoshi, T., 2000. Stochastic DEA for restructure strategy: an application to a Japanese petroleum company. Omega 28 (4), 385–398.
Talluri, S., Narasimhan, R., Nair, A., 2006. Vendor performance with supply risk: a chance-constrained DEA approach. Int. J. Prod. Econ. 100 (2), 212–222.
Thore, S., 1987. Chance-constrained activity analysis. Eur. J. Oper. Res. 30 (3), 267–269.
Torgo, L., 2011. Data Mining with R: Learning with Case Studies. CRC Press, Boca Raton.
Tsolas, I.E., Charles, V., 2015. Incorporating risk into bank efficiency: a satisficing DEA approach to assess the Greek banking crisis. Expert Syst. Appl. 42, 3491–3500.
Udhayakumar, A., Charles, V., Kumar, M., 2011. Stochastic simulation based genetic algorithm for chance constrained data envelopment analysis problems. Omega 39 (4), 387–397.
Vapnik, V., 1995. The Nature of Statistical Learning Theory. Springer, New York.

Vapnik, V., 1998. Statistical Learning Theory. John Wiley & Sons, New York.
Wang, Z., Feng, C., 2015. Sources of production inefficiency and productivity growth in China: a global data envelopment analysis. Energy Econ. 49, 380–389.
Wang, W.K., Lu, W.M., Lin, Y.L., 2012. Does corporate governance play an important role in BHC performance? Evidence from the US. Econ. Model. 29 (3), 751–760.
Wanke, P.F., 2012a. Capacity shortfall and efficiency determinants in Brazilian airports: evidence from bootstrapped DEA estimates. Socio-Econ. Plan. Sci. 46 (3), 216–229.
Wanke, P.F., 2012b. Efficiency of Brazil's airports: evidences from bootstrapped DEA and FDH estimates. J. Air Transp. Manag. 23, 47–53.
Wanke, P., Barros, C., Faria, J.R., 2015a. Financial distress drivers in Brazilian banks: a dynamic slacks approach. Eur. J. Oper. Res. 240 (1), 258–268. http://dx.doi.org/10.1016/j.ejor.2014.06.044.
Wanke, P., Barros, C., Macanda, N.P.J., 2015b. Predicting efficiency in angolan banks: a two-stage TOPSIS and neural networks approach. S. Afr. J. Econ. http://dx.doi.org/10.1111/saje.12103.
Wanke, P., Barros, C.P., Nwaogbe, O.R., 2016a. Assessing productive efficiency in Nigerian airports using Fuzzy-DEA. Transp. Policy 49, 9–19.
Wanke, P., Azad, M.D., Barros, C.P., 2016b. Predicting efficiency in Malaysian Islamic banks: a two-stage TOPSIS and neural networks approach. Res. Int. Bus. Financ. 36, 485–498.
Wolfe, P., 1961. A duality theorem for non-linear programming. Q. Appl. Math. 19, 239–244.
Yang, N., Wen, F.S., 2005. A chance constrained programming approach to transmission system expansion planning. Electr. Power Syst. Res. 75 (2/3), 171–177.
Yang, N., Yu, C.W., Wen, F., Chung, C.Y., 2007. An investigation of reactive power planning based on chance constrained programming. Int. J. Electr. Power Energy Syst. 29 (9), 650–656.
Yu, M.M., 2010. Assessment of airport performance using the SBM-NDEA model. Omega 38 (6), 440–452.
Zhu, J., 2003. Quantitative Models for Performance Evaluation and Benchmarking: Data Envelopment Analysis with Spreadsheets and DEA Excel Solver. Springer, New York.