Energy Economics 45 (2014) 111–119

Contents lists available at ScienceDirect

Energy Economics

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

Design of experiments applied to environmental variables analysis in electricity utilities efficiency: The Brazilian case

Rafael C. Leme [a], Anderson P. Paiva [a], Paulo E. Steele Santos [b], Pedro P. Balestrassi [a,*], Leandro de Lima Galvão [b]

[a] Institute of Industrial Engineering and Management, Federal University of Itajuba, Av. BPS 1303, Itajubá, MG, Brazil
[b] TR Soluções, INCIT, Av. BPS 1303, Itajubá, MG, Brazil

ARTICLE INFO

Article history:
Received 1 October 2013
Received in revised form 19 March 2014
Accepted 19 June 2014
Available online 5 July 2014

JEL Classification:
L94 Electric Utilities
L97 Utilities: General
C15 Statistical Simulation Methods: General
C44 Operations Research • Statistical Decision Theory
C52 Model Evaluation, Validation, and Selection
C9 Design of Experiments

Keywords:
Efficiency Analysis
Environmental variables
Electricity utility
Data envelopment analysis
Design of experiments

ABSTRACT

Benchmarking plays a central role in the regulatory scene. Regulators set tariffs according to a performance standard and, if the companies can outperform such a standard, they can retain the gains observed by such outperformance. Efficiency performance is usually assessed by comparison (or a benchmark) against either other companies or the company's own historical performance. This paper discusses the impact of environmental variables on the efficiency performance of electricity distribution companies. Indeed, such variables, which are argued to be unmanageable, may affect the electricity utilities' performance. Thus, this paper proposes a simulation methodology based on design of experiment philosophy for statistically testing environmental variables and the interactions among them, enabling regulators to build the best suited semi-parametric two-stage model of electricity utility benchmarking analysis. To demonstrate the power of the proposed approach, experimental simulations are carried out using real data published by Brazil's regulator. The results show that environmental variables may impact efficiency performance linearly and nonlinearly.

1. Introduction

When competition is feasible, the market determines the optimal combination of price and quality. A market automatically converges, economists believe, to prices that reflect efficient costs for suitable products. In natural monopolies, however, which are what electricity distribution networks are, competition is not feasible. Here optimal cost efficiency and quality standards are determined by the hand of the regulator.

Initially, the regulator makes sure that a utility's costs are all recovered and supplemented by a reasonable return. This approach, known as rate-of-return, sets prices based on observed costs; it offers a company no stimulus and no incentive, to be efficient. It is the regulator's duty, however, to not only ensure non-discriminatory access charges but to also design charges that reflect efficient costs.

Thus, regulators around the world have investigated and implemented numerous efficiency-based regulatory approaches by benchmarking electricity utilities (see Jamasb and Pollitt, 2001). Benchmarking analysis is usually integrated into a *RPI-X* mechanism, a mechanism that refers to incentive regulation (Brophy Haney and Pollitt, 2009; Evans and Guthrie, 2006; Joskow, 2008). Such a regulation model adjusts tariffs based on the rate of inflation — the Retail Price Index (*RPI*) — and on the Distribution System Operators' (DSOs) efficiency compared to that of the reference company, and efficiency is known as the expected efficiency savings, or *X-factors*.

The most widely used of the benchmark techniques are Corrected Ordinary Least Squares (COLS), Stochastic Frontier Analysis (SFA), and Data Envelopment Analysis (DEA). The first two select an equation that defines the relationship between the explanatory and dependent variables. As dependent variables are usually considered the costs and the explanatory variables are usually the services to be analyzed. Such approaches are defined as parametric. Then, the efficiency gap to be closed is defined by the error between the selected equation and the actual value of costs. The third benchmark technique, DEA, was first proposed by Charnes et al. (1978), and it is a non-parametric technique. DEA uses linear programming to define an envelope around observations. The frontier is defined by the envelope that efficient firms form around less efficient ones. The distance separating the two defines the efficiency gap. Shuttleworth (2005) pointed out that DEA used on

* Corresponding author. Tel.: +55 35 3629 1161.
E-mail address: pedro@unifei.edu.br (P.P. Balestrassi).

http://dx.doi.org/10.1016/j.eneco.2014.06.017
0140-9883/

electricity distribution business has limitations regarding the small sample size and statistical inferences. Jamasb and Pollitt (2003), however, have shown that to deal with the sample size problem regulators can use cross-country benchmarking. These limitations may also be overcome using bootstrap techniques (see Odeck, 2009), robust solutions of linear programming problems contaminated with uncertain data (see Sadjadi and Omrani, 2008), or Bayesian inferences (see Tsionas and Papadakis, 2010).

Nevertheless, other issues arise when DSO results are influenced by certain contextual factors. In the utility business, these contextual factors are also known as environmental variables. Businessmen argue that unmanageable environmental variables usually affect a firm's costs and quality efficiency performance. Weather conditions, for example, have been shown to be highly correlated to the reliability of a distribution network (Billinton and Allan, 1984; Coelho et al., 2003; Domijan et al., 2003; Wang and Billington, 2002). It may be argued, on the other hand, that utilities adapt their operations and investment to mitigate the environment's adverse effects. Indeed, Yu et al. (2009), who studied the effects of weather conditions in the cost and quality performance of UK DSOs, concluded that on average the impact is small. The authors argued that considering one or another output in the efficiency analysis might internalize the effect of the contextual factors.

In most countries, contextual factors are summed up as weather conditions. However, in an expansive country like Brazil, the environmental variables may also include the regions' disparate conditions such as salary and population density. Indeed, as further discussed, besides a wide area country, Brazil has a large heterogeneity that characterizes the Brazilian utility industry. For instance, some DSOs are responsible for just a single town with high consumption density, whereas others are responsible to deliver energy in large states. Other discrepancies are discussed in a recent study by Brazil's regulator, who suggested that four environmental variables could explain DSOs' inefficiency (ANEEL, 2010, 2011), as further discussed.

In analyzing environmental variables, two drawbacks arise: first, one must use a mathematical model which measures the influence of environmental variables on efficiency score with a reasonable degree of confidence; second, one must choose a set of environmental variable and test their significance on the mathematical model as well as the interactions among them.

For the first drawback, the semi-parametric two-stage approach has been gaining great attention in the literature (Chilingerian and Sherman, 2004; Ray, 2004; Ruggiero, 2004). Such approach regresses the efficiency estimated using traditional DEA (first stage) on environmental variables (second stage), which is also known as two-stage DEA. In exploring semi-parametric two-stage approach, Simar and Wilson (2007) thus proposed a data-generating process consistent with non-parametric efficiency estimates, avoiding problems with serially correlated data. In the approach proposed by Simar and Wilson (2007), the bootstrap technique makes model inference possible and feasible. Such approach is used in this paper to model the environmental variables' impact in efficiency scores. As for the second drawback, despite several applications in the literature, no straightforward way exists to identify regressors; the applications require different statistical techniques and often end up being a process of trial and error.

In this sense, this paper presents an alternative approach based on design of experiments (DOE). DOE is a collection of statistical techniques capable of generating and analyzing experimental designs in which several factors are varied simultaneously. These experimental design methods, introduced by Fisher (1966), are widely used in production and operation management, as well as manufacturing and quality control (Montgomery, 2009). The methodology seeks to plan an experimental design so that appropriate data can be obtained that lead to good conclusions, by using a minimal number of experiment runs. Thus, in analyzing environmental variables, the DOE technique may help evaluate the significance of each factor in estimating efficiency, as well as it is also useful in assessing the statistical significance of interaction among variables. The regulator is then able to assess the impact of environmental variables on DSO inefficiency by using the two-stage DEA model.

Recently, DOE has been used extensively in applications related to simulation analysis (see Kleijnen, 2005). Balestrassi et al. (2009), for instance, used the fractional and full factorial designs to better determine the parameters of an artificial neural network (ANN), bypassing the trial-and-error technique. Sun and Li (2013), in turn, have used designed simulation experiments in optimizing operating room scheduling. In the same way, Oliveira et al. (2011) presented a novel approach, using mixture design of experiments, to adjust the conditional value at risk metric for a mix of contracts on the energy market.

Importantly, this paper has no ambition of constructing a new benchmark model. It intends rather to propose, with the help of DOE, a simulation approach for environmental variable analysis. Its major contribution resides in the proposing of a simulation methodology for statistically testing environmental variables and the interactions among them, enabling analysts to build the best suited second stage model to electricity utility benchmarking analysis, when two-stage semi-parametric approach is considered. To demonstrate the power of the proposed approach, experimental simulations are carried out using real data published by Brazil's regulator. The results show that environmental variables may impact efficiency performance linearly or nonlinearly, so that the approach proposed in this paper may avoid a misspecified model.

The remainder of this paper is organized as follows: Section 2 briefly reviews some incentive regulation concept and Brazil's network regulation background; Section 3 presents the semi-parametric two-stage DEA used in this paper; Section 4 discusses the features of DOE philosophy and lays out the methodology adopted; and Section 5 presents the results and a discussion of the experimental simulated dataset. Finally, Section 6 presents our conclusions.

2. Regulation of electricity distribution network in Brazil

When dealing with utility regulation, regulators may consider two incentive mechanisms: revenue cap and price cap. The revenue cap limits the predetermined level of annual revenue that a DSO can collect from its consumers. The price cap defines, based on the prices of different products (access, energy, demand, etc.), the annual permitted revenue. In these regulatory scenarios, prices are decoupled from observed costs.

The annual permitted revenue in incentive regulation consists of setting tariffs according to a performance standard and, if the companies can outperform such a standard, they can retain the gains observed by such outperformance. Hence, the DSO is motivated to operate efficiently and thereby cut down on costs to increase the shareholders' profits. Under such mechanism, benchmarking plays a central role. In fact, regulators' measure of the performance standard is usually assessed by comparing it against the company's own historical performance and/or against other companies. Thus, based on each DSO's performance analysis, the regulator sets its initial annual permitted revenue P_0. Once the permitted income is so defined, the cost allocation establishes how to collect it from the end-users (Steele Santos et al., 2012).

Nevertheless, when a capital-intensive industry (like that of an electricity utility) adopts an efficiency-based regulatory approach it can lead to a degradation in service quality, so one must consider the *X-factors*. By establishing *X-factors the* regulator considers a benchmark model for efficiency and quality; it uses *X-factors* to close any eventual inefficiency gaps. The benchmark is usually set by the identification of the most efficient practice in the sector. Jamasb and Pollitt (2003) presented some approaches for efficiency benchmark models. Also, Ajodhia and Hakvoort (2005) discussed quality regulation approaches.

Eq. (1) describes the usual framework of the incentive regulation mechanism: each company faced with a permitted annual revenue P_0, which is kept for the subsequent period, say a year, adjusted only to a retail price index *RPI* less than the efficiency parameter *X*.

$$P_1 = P_0(RPI - X) \tag{1}$$

Eq. (1) is applied for all subsequent periods within the same regulatory lag (usually four or five years). Once the regulatory lag is over, a price review takes place, defining a new P_0, and the process starts all over again.

Note that, although the efficiency parameter *X* explicitly accounts for an efficiency adjustment along the regulatory period, efficiency may also be promoted when the regulator sets the initial permitted revenue P_0. Therefore, in addition to the problem of estimating the efficiency standard, the regulator may also decide how to translate it into tariffs.

Brazil has 63 DSOs, with each having a franchise over a specific territory. The federal government gives directives through concession contracts, which are standard for all companies. The institution responsible for overseeing these contracts is Brazil's federal regulator, the Brazilian Electricity Regulator Agency,[1] an independent entity empowered with setting tariffs, quality standards, and penalties.

Brazil's electricity utility industry is a study in diversity. Some companies deliver energy in the large states of the Amazon region; other companies are responsible for just a single town with high consumption density. Indeed, states in the south are wealthier and more densely populated than other regions and thus tend to administer to more complex franchise areas.

The distribution network is formally defined as every installation below 230 kV, including transformers whose secondary voltage is below this value. Tariffs for DSOs are set according to each company's franchise contract but based on common standards. The franchise contract establishes a price-cap with tariffs adjusted yearly for *RPI-X* with a regulatory period varying between three and five years.

The regulated costs are categorized as controllable and pass-through costs. The manageable costs include all distribution costs: capital expenditure, operational and maintenance costs. Network losses in the distribution network, although regulated as manageable, are defined as pass-through. At the national level, the manageable costs of distribution represent about 29% of the electricity bill.

Since 2012, a new price-setting methodology has been in place. The current approach mainly innovates at setting the operating costs in Brazil. Following what has been done in other countries around the world, the Brazilian regulator has been using benchmark analysis to set *X-factors*. In Brazil, *X-factors* is designed to capture three different inefficiency gaps within regulatory lag: P_d seeks to measure the DSO's average productivity, *Q* refers to the quality standard and *T* measures eventual inefficiency gap in operational and maintenance costs. Thus, Brazil's *X-factor* is defined as:

$$X = P_d + Q + T. \tag{2}$$

P_d is obtained by the total productivity considering capital expenditure and operational expenditure costs, whereas the factor *Q* is set by minimal quality standard. As for factor *T*, which is the focus of this paper, it is related to efficiency score, estimated by benchmarking operating costs the operating costs in order to measure relative inefficiencies. Brazil's methodology estimates efficiency on operational and maintenance costs by using semi-parametric two-stage approach in order to set *T-factors*, which seek to close eventual inefficiency gap on operating costs. First, the efficiency scores are calculated by DEA considering DSO's usual operating variables and non-decreasing return to scale. A COLS is also applied to the same data set to verify the consistency of results. In the second stage, the efficiency scores are normalized according to environmental variables to make provision for heterogeneity among the regulated companies. Such environmental variables are chosen to describe the wide discrepancy that characterizes the Brazilian utility industry, as further discussed. The semi-parametric two-stage model used is set based in basic statistical tests in the second stage analysis, ignoring potential interactions among environmental variables. Moreover, the analysis is carried out by splitting DSOs into two groups: big companies (which delivered more than 1 TWh) and small companies (otherwise). This is so, since companies with very different size tend to present different return to scales.

In order to extend the analysis reported by the Brazilian regulator, the next section presents the two-stage DEA approach which may be used in environmental variable analysis. Nonetheless, the approach proposed here may easily be extended to other two-stage benchmark models.

3. Two-stage DEA applied to electricity distribution sector

Data envelopment analysis is a non-parametric technique that computes efficiency. The idea is to find, using linear programming, the best practices from a sample of firms within a set of comparable Decision Making Units (DMUs). A feasible production set is generally defined by a convex region containing sample observations of firms' inputs and outputs. The relative performance of DMUs is obtained according to their location within a production set. The Pareto–Koopmans-efficient firms, the ones that are fully efficient,[2] then delineate an efficiency isoquant frontier by the envelope they form around the less efficient firms. A company's efficiency is determined by the distance measure of a DMU from the isoquant.

Efficiency can be measured in many ways. Farrell (1957) proposed a radial measure. It may be used either in the input or output spaces. Using radial distance, the efficiency score defines the amount that DMU must reduce (increase) its input (output) to lie on the isoquant. Thus, the radial distance to the frontier is considered a measure of inefficiency; i.e., it defines the inefficiency gap. In other words, the DEA sets the benchmark DMUs and the distances of other firms to them. When the mix within inputs and outputs in movements towards the frontier is not preserved, one has a non-radial efficiency score (Cooper et al., 2007; Thanassoulis et al., 2008). This paper, however, does not consider such an approach.

A number of DEA models have been developed for different purposes. Basically, DEA can be input (or output) oriented. The input computes the minimal resources required to produce a given level of outputs. The output maximizes the production for a given level of inputs. Depending on the case, one may be interested in reducing inputs and increasing outputs altogether. In these cases, non-oriented DEA models may be used.

Furthermore, DEA can be specified as Constant Returns to Scale (CRS) or Variable Returns to Scale (VRS). The CRS approach considers that an increase in the inputs leads an increase in the output by the same proportion. Thus, the relative efficiency is unaffected by the size of the company. VRS, on the other hand, is defined when the linearity on input/output fails to hold. In this case, companies are more likely to be compared by size. When the returns to scale are not constant, non-decreasing (NDRS) and non-increasing (NIRS) returns to scale may also be defined.

When environmental variables affect a company's performance, one may use the two-stage semi-parametric approach known as two-stage DEA. The first stage is defined by traditional DEA approaches. In the second stage, the estimated efficiency is regressed on environmental variables, which are considered unmanageable. Then, the efficiency score may be corrected considering the contextual factors. The problem in using two-stage DEA is that the results may be biased, since environmental variables are highly correlated and efficiency scores do not follow normal distribution. In handling these problems, one may

[1] http://www.aneel.gov.br/?idiomaAtual=1.

[2] Cooper et al. (2007, p.45) state that "A DMU is fully efficient if and only if it is not possible to improve any input or output without worsening some other inputs or outputs."

apply the bootstrap technique, such as the one proposed by Simar and Wilson (2007).

In this paper, we use the two-stage DEA approach as Brazil's regulator does, an approach similar to that proposed by Simar and Wilson (2007). Here the bootstrap technique is used to make inferences on the second stage regression. The model is described below.

3.1. First stage DEA

In regulatory contexts, DSOs must provide services to all consumers in their service area considering the operational and maintenance costs allowed by the regulator (Jamasb and Pollitt, 2001; Lowry and Getachew, 2009). Thus in the first stage of DEA, input may be considered as the operational and capital costs. Furthermore, the most commonly used outputs regard the number of costumers (NC) and energy delivered (ED). The former represents commercial costs and the latter represents network density. Besides these outputs, the network length (NL) is also considered. NL represents maintenance costs.

The regulator expects DSOs to use their permitted annual income efficiently. If they fail to, the regulator cuts off the regulatory incomes as to stimulate DSO best practices. Thus, in analyzing DSO performance, the regulator usually considers the input-oriented DEA approach. Note that in this model, only one input is considered. Thus, Farrell's (1957) radial measure of efficiency is used, with no loss of generality. The DEA model used in this paper is given as follows:

$$\min_{\theta_i,\lambda}\left\{\theta_i\left|\begin{array}{c} Y\cdot\lambda\geq y_i \\ X\cdot\lambda\leq\theta_i x_i \\ \lambda\geq 0 \\ \lambda^T u\geq 1\end{array}\right.\right\}. \tag{3}$$

In Eq. (3), θ_i is the efficiency parameter for the i-th DMU; X and Y are the $n\times m$ input and $n\times r$ output matrix, respectively, and λ is the weight parameter. Furthermore, n, m, and r are the amount of DMUs, inputs, and outputs, respectively. The vector u is a column vector with all entries equal to one. Finally, x_i and y_i are the input and output vectors of the DMU analyzed.

The DEA model of Eq. (3) defines the NDRS approach since $\lambda^T u\geq 1$. If θ_i reaches one, the input level is optimized and cannot be reduced for the given level of output. On the other hand, if θ_i is less than one, a cut off should be made to DSO costs to reach the given level of outputs. Thus the efficiency frontier is defined by θ_i equaling to one.

Indeed, this is the approach used by the Brazilian regulator, as discussed in Section 2. However, the approach described in this paper may be applied with any DEA approach, as well as any efficiency analysis.

3.2. Second stage DEA

In the first stage, the technical efficiency is obtained, i.e., the ability of companies to minimize inputs to produce a given level of outputs. In the second stage, however, one is interested in the impact of contextual factors. Thus, the two-stage DEA allows environmental variables to be considered in the efficiency analysis. However, as stated above, usually, the first-stage variables are highly correlated to second-stage variables, so the regression results are likely to be biased. To overcome this drawback, Simar and Wilson (2007) proposed using the bootstrap technique to make statistical inferences. Thus, if a company must confront environmental variables z_i, one may write the regression:

$$\delta_i=\psi(z_i,\beta)+\varepsilon_i\geq 1. \tag{4}$$

In Eq. (4), ψ is a general function of β, usually defined as linear or polynomial relationship on z. Furthermore, ε_i is the error independent and identically normally distributed. Finally, $\delta_i=1/\theta_i$.[3] To solve this problem then, one may use censored regression. Censored regression is a regression model where efficient DMUs are censored at 1.00. It estimates the relationship between the explanatory variable and the censored dependent variables.

The bootstrap technique for solving a two-stage DEA is as follows:

I. Compute the first stage efficiency using Eq. (3).
II. With the help of regression censored in 1.00 by left, estimate $\hat{\beta}$ of β, and the error variance $\hat{\sigma}_\varepsilon^2$ of Eq. (4)'s σ_ε^2, considering only the inefficient companies.
III. Loop over (a) and (c) L times to obtain the bootstrap estimates of β and σ_ε^2;
   a. For each inefficient firm, draw ε_i^* as from the normal distribution $N\left(0,\hat{\sigma}_\varepsilon^2\right)$, with left truncation at $\left(1-z\cdot\hat{\beta}\right)$.
   b. Compute efficiencies $\delta_i^*=z_i\cdot\hat{\beta}+\varepsilon_i^*$.
   c. With the help of regression censored in 1.00 by left, estimate $\hat{\beta}^*$ of β, and the error variance $\hat{\sigma}_\varepsilon^{2*}$ of σ_ε^2.
IV. Use the bootstrap values to correct the estimate $\hat{\beta}$ of β, and the error variance $\hat{\sigma}_\varepsilon^2$ of σ_ε^2, and construct estimated confidence intervals.

Once estimated, the impact of environmental variables may be adjusted for each company, considering, for example, the mean scenario for these variables. However, prior to this adjustment, the model analysis and selection may be performed, so that the chosen are only variables that truly impact the efficiency performance. A widely used approach in model selection is the consistent information criteria index, as the Bayesian information criterion (*BIC*) (Schwarz, 1978), which is given in Eq. (5).

$$BIC=-2\cdot\ell+\ln(n)\cdot\sum_{i=1}^{p} i \tag{5}$$

In Eq. (5), ℓ refers to the likelihood of estimated model; p stands for the number of parameters estimated in regression analysis, and n is the number of observed outcomes.

In general, when using information criteria, some candidates are selected and estimated. Thus, one may compute *BIC* using Eq. (5) for each estimated model. As a rule of thumb, one chooses the model that yields the smallest value.

4. Design of experiments methodology for assessment of environmental variables

In model analysis, one may use multivariate tools as factor analysis and principal component analysis. The former seeks to extract composite factors from a set of observed variables; the latter uses the covariance information matrix to rotate a variable's reference axis to identify how the variance of data is explained. Such techniques, however, usually simplify the analyzed model based on eigenvalue analysis. An alternative approach lies in the design of experiment (DOE) methodology.

DOE, with its numerous success stories (see Montgomery, 2009), is considered one of the most important methodologies for researchers handling experiments in practical applications (Balestrassi et al., 2009). Today, DOE resources are incorporated in many statistical software packages that ease calculation and interpretation of results (Chan and Spedding, 2001).

According to Montgomery (2009), DOE is a collection of statistical techniques capable of generating and analyzing experimental designs in which several factors are varied at once rather than one at a time. Among the most common available designs are the screening designs,

[3] $\delta_i=1/\theta_i$ may vary from 1 to ∞, simplifying regression model, since it is only censored by left.

Table 1
Full factorial design for *F* environmental variables.

	r	z_1	z_2	z_3	z_4	z_5	… z_i	*BIC*
	+1	−1	−1	−1	−1	−1		
F = *1*	2	+1	−1	−1	−1	−1		
	3	−1	+1	−1	−1	−1		
F = *2*	4	+1	+1	−1	−1	−1		
	5	−1	−1	+1	−1	−1		
	6	+1	−1	+1	−1	−1		
	7	−1	+1	+1	−1	−1		
F = *3*	8	+1	+1	+1	−1	−1		
	9	−1	−1	−1	+1	−1		
	10	+1	−1	−1	1	−1		
	11	−1	+1	−1	+1	−1		
	12	+1	+1	−1	+1	−1		
	13	−1	−1	+1	+1	−1		
	14	+1	−1	+1	+1	−1		
	15	−1	+1	+1	+1	−1		
F = *4*	16	+1	+1	+1	+1	−1		
	17	−1	−1	−1	−1	+1		
	18	+1	−1	−1	−1	+1		
	19	−1	+1	−1	−1	+1		
	20	+1	+1	−1	−1	+1		
	21	−1	−1	+1	−1	+1		
	22	+1	−1	+1	−1	+1		
	23	−1	+1	+1	−1	+1		
	24	+1	+1	+1	−1	+1		
	25	−1	−1	−1	+1	+1		
	26	+1	−1	−1	+1	+1		
	27	−1	+1	−1	+1	+1		
	28	+1	+1	−1	+1	+1		
	29	−1	−1	+1	+1	+1		
	30	+1	−1	+1	+1	+1		
	31	−1	+1	+1	+1	+1		
F = *5*	32	+1	+1	+1	+1	+1		
	.							
	.							
	.							
F	2^F							

fractional or full factorial designs, response surface methodology, evolutionary operation, and mixture designs.

The advantage in using simulated experimental designs is that one is able to gather information, while avoiding trial and error analysis, about a process through systematic planned experiments. In a decision model approach where the analyst must choose the presence or absence of a variable in a model, one has 2^p candidates to be evaluated.

The use of DOE in analysis of simulated experiments was reviewed by Kleijnen (2005). DOE is used in simulation for sensitivity analysis of the considered factors. Using statistical analysis, one may detect the unimportant factors and thereby simplify the simulation model. An important thing to remember when using DOE for simulation is that the main goal is not optimization itself. Efforts are dedicated to find robust policies or decisions.

In this paper, we propose using DOE to gather information, by simulating a two-level factorial design, about the degree of influence of environmental variables on an efficiency score. In this kind of experimental design, researchers examine how the presence (level +1) or the absence (level-1) influences each environmental variable in DEA model. Considering *F* environmental variables that may impact the efficiency score, a full factorial design is presented in Table 1.

In analyzing the simulated results, the metamodel considering all *F* main effects and iterations may be used. This metamodel is shown in Eq. (6).

$$BIC_r = \beta_0 + \sum_{i=1}^{F} \beta_i z_i + \sum_{i,j=1,i<j}^{F} \beta_{i,j} z_i z_i + \ldots + \beta_{1,2,\ldots,F} z_1 z_2 \cdots z_F + \varepsilon \qquad (6)$$

In Eq. (6), β_0, β_i, $\beta_{i,j}$, …, $\beta_{1,2,\ldots,F}$ are known as effect coefficients. Furthermore, z_i is the observed value of environmental variable *i*, in the same way of Eq. (4). Finally, *r* is the *r*-th run of experimental design. Note that the experimental design of Table 1 is an orthogonal design, so that the estimates of Eq. (6) are not correlated.

The metamodel of Eq. (6) may be obtained with the help of ordinary least squares (OLS) or maximum likelihood estimation (MLE). The error ε is considered independent and identically normally distributed. Once estimated, the metamodel may be then analyzed using *BIC* and two sample-t hypotheses and nonparametric tests. The aim is to verify the difference in mean response between the two levels of each factor, as well as the interactions among them. The critical value for the statistics used is obtained using a 5% significance level.

The two-stage DEA approach described in Section 3 is used in the comparative efficiency analysis, where the choice of variables and model specification is a fundamental aspect of the DEA benchmarking. Recall, however, that this paper aims to propose a simulation approach to help analysts analyze the environmental impact on efficiency performance. Thus, no methodological ambitions exist for defining which input or output variables ought to be considered; the goal is to focus on a DOE based simulation approach to analyze the environmental variables.[4] Thus, we follow the model of Brazil's regulator, discussed briefly in Section 2. Hence, in the input-oriented DEA model laid out in Section 3.1, the input is the accounting operational expenditure (Opex) and the outputs are the network length (NL), number of consumers (NC), and energy delivered (ED). In its benchmarking analysis, Brazil's regulator chooses these three output variables as cost drivers, arguing that such an analysis is the most consistently adopted by regulators of other countries. Remarkably, under Brazilian regulation, capital costs are treated separately. Following DEA benchmarking, to analyze the impact of context variables, one may define the environmental variables. In small countries, the number of contextual variables, if they exist at all, may be few. In an expansive country, environmental variables are more likely to abound. In the UK, for example, the DSO performance is affected by weather variables only (see Yu et al., 2009). To explain inefficiency performance around Brazil, in contrast, Brazil's regulator publishes a technical report that mentions four environmental variables (ANEEL, 2010, 2011): mean wage, precipitation index, complexity index, and consumer density. Indeed, this group of environmental variables is meant to account for the heterogeneity among the regulated companies. Again, this paper has no ambition for defining which environmental variables ought to be considered, but to use simulation based on DOE methodology for statistically test environmental variables and the interactions among them, enabling the regulator to build the best suited second stage model to electricity utilities benchmarking analysis, when considering the semi-parametric two-stage approach.

Thus, we first estimate DSOs' efficiency using DEA as described above. Once the first stage is achieved, the environmental variables are regressed against the efficiency score obtained in the first stage considering full factorial analysis discussed above. For this sake, the censored regression is used jointly with bootstrap, as discussed in Section 3.2. The results are then analyzed and a discussion about the statistical significance of environmental variables in the efficiency scores of DSOs is presented. The decision analysis is based on *BIC* in order to discuss model specification, and two sample-*t* tests are performed for each factor, identifying the most important (environmental) factors for inefficiency, as well as their interactions. Finally, the results are compared to Brazil's regulator model, explicating the difference between them.

[4] Indeed, surveys presented in literatures shows that there is no consensus on the best choice of variables to measure cost efficiency of distribution networks (Ajodhia, 2006; Jamasb and Pollitt, 2001).

Table 2
Input and outputs variables summary statistics.

Variable	n	Mean	Min	Max
Opex ($)	203	397.330.519	75.600.376	1.869.287.128
ED (MWh)	203	4.229.024	685.523	18.053.158
NC	203	1.923.808	434.378	6.832.546
NL (km)	203	82.909	8.040	460.219

Table 3
Environmental variables summary statistics.

Variable	n	Mean	Min	Max
MS	203	8.0779	7.5631	8.6835
PI	203	7.0019	0.0000	7.8038
CI	203	0.2145	0.0284	0.4581
CA	203	9.9693	9.0137	10.7920

MS, PI and CA are natural logarithm of the original data

5. Data set and experimental simulation

The two-stage DEA benchmarking analysis presented in this paper uses the data set published by Brazil's regulator for the proposal of operational costs based on data for the 61 main distribution companies (from 2003 to 2009). Such data may be found on the regulator's website[5] and is available as electronic supplementary material to this paper. Since less than 10% of the energy delivered in Brazil comes from small companies, in this paper, we analyze only the big companies group, yielding 29 DSOs. Table 2 presents descriptive statistics for input and outputs variables.

Note, from Table 2, that Opex presents a wide range. The DSO with minimum allowed Opex is permitted to collect about $ 75.6 million annually to cover its operational and maintenance costs, whereas the DSO with maximum allowed Opex can collect more than $ 1.8 billion. Thus, even small difference in efficiency score may represent a great amount of money overpaid/underpaid by the consumers. If consumers overpay the operating costs, DSOs earn more than the fair Opex. On the other hand, if the consumers underpay the operating costs, DSOs may have a deficit in its account.

Thus, in showing the power of the proposed approach in this paper, we first describe the actual Brazilian second stage model. Next, the DOE simulation approach is performed to identify the best suited second stage model. The comparison between the second stage model obtained by DOE philosophy and the model used by Brazilian Regulator is then presented.

5.1. The second stage model in Brazil

Brazil's regulator has, as described above, established four contextual factors that may impact DMUs' performance. The first environmental variable considered refers to mean salary (MS), which represents the mean employees incomes based in a government database. Such environmental variables tend to capture the wide variation on cost of living among the regions, representing the richness of each region. The second environmental variable regards a precipitation index (PI), which is agreed to be correlated to operational and maintenance costs. Brazil's regulator advocates that, depending on weather condition, DSOs may face an increase in outages, which demands more operational and maintenance teams. A complexity index (CI) is defined as the third environmental variable. The CI measures difficulty faced by DSOs in combating non-technical losses, such as electricity theft. A company that has a concession area with a higher degree of complexity usually has greater difficulty in reducing their losses compared to the other whose concession area has a lower index, so it may involve different operating costs. Finally, the fourth environmental variable is consumer per area (CA), which captures consumer density. Indeed, DSOs that have a high concentration of consumers tend to observe lower operating and maintenance costs.

These environmental variables are computed based on different approaches proposed by Brazil's regulator. For instance, MS and CA are, respectively, a simple average and ratio based on government and regulator databases' data, respectively; CI is based on an econometric approach that refers to socioeconomic aspects such as electricity theft; PI is obtained based on isohyets curves on the DSO area. Table 3 summarizes the statistics on these environmental variables.

Based on the observation of these environmental variables, Brazil's regulator proposes the use of a linear model to represent the relationship among efficiency score obtained in the first stage and environmental variables. Such model is given by:

$$\delta_i = -4.45 + 0.77z_{MS} + 0.27z_{PI} + 0.46z_{CI} - 0.22z_{CA} + \varepsilon. \tag{7}$$

In Eq. (7), z_{MS}, z_{PI}, z_{CI} and z_{CA} refer to the environmental variables MS, PI, CI and CA, respectively. Importantly, all these environmental variables have shown statistically significance for the linear model of efficiency score.[6] The index criteria of the model of equation is $BIC = 1052.9$. The econometric model of Eq. (7) is further used to compare Brazil's regulator model to the one obtained by DOE simulation approach, which is discussed next.

5.2. Second stage model by the means of DOE

With the environmental variables described above, DOE is now used for sensitivity analysis of environmental variables. For this sake, two replicates of full factorial design are used, yielding 32 simulated experiments.[7] The approach used is given by Eq. (4), with linear and interaction relationships being considered. The presence of each environmental variable in a model may be evaluated, so that the evaluations consider the presence (+1) or absence (−1) of each contextual factor. In the model estimation, the bootstrap technique discussed in Section 3.2 is used, through the full factorial design, to obtain regression-coefficient distribution.

The analysis on simulation results may be carried out using a metamodel similar to Eq. (6). Recall that the response regards *BIC*, which is a consistent criteria in regression model selection. As for the used metamodel, we truncated in two-factor interaction, yielding the metamodel of Eq. (8). The simulated experiments are presented in Table 4.

$$BIC_r = \beta_0 + \sum_{i=1}^{F}\beta_i z_i + \sum_{i,j=1,i<j}^{F}\beta_{i,j} z_i z_i + \varepsilon \tag{8}$$

From the simulated experimental results of Table 4, one can plot the main effects for *BIC*, as illustrated in Fig. 1. In this graph, one can see how each factor impacts the response variable, i.e., how the presence and absence of each environmental variable improves/worsens *BIC* specification.

In Fig. 1, for instance, one can see that the environmental variables MS and PI significantly impact the *BIC*. It may be observed by noting that when environmental variables, say MS, switch from absence (−1) to presence (+1) model analysis, the *BIC* reduces from about 1800 to less than 1200, improving *BIC* specification.[8] Interestingly, CI seems to have no influence on the *BIC* by switching from absence to presence, whereas CA seems to slightly worsen *BIC* specifications.

[5] http://www.aneel.gov.br/?idiomaAtual=1.

[6] The significance level used by Brazillian regulator is 10%.

[7] Two replicates are chosen, since the simulated model is based on bootstrap approach, so that replicating factorial designs yields more robust model analysis.

[8] Recall that selection based on BIC chooses according “smaller is better” criteria.

Table 4
2^4 full factorial design for 4 environmental variables.

Regression	MS	PI	CI	CA	BIC	
1	−1	−1	−1	−1	2257.07	2257.07
2	+1	−1	−1	−1	1189.78	1187.33
3	−1	+1	−1	−1	1286.27	1275.79
4	+1	+1	−1	−1	1112.07	1107.76
5	−1	−1	+1	−1	2261.57	2261.59
6	+1	−1	+1	−1	1157.07	1174.92
7	−1	+1	+1	−1	1353.93	1352.45
8	+1	+1	+1	−1	1083.44	1083.11
9	−1	−1	−1	+1	2258.72	2258.55
10	+1	−1	−1	+1	1114.64	1107.22
11	−1	+1	−1	+1	1287.38	1282.83
12	+1	+1	−1	+1	1085.00	1091.64
13	−1	−1	+1	+1	2261.49	2261.46
14	+1	−1	+1	+1	1089.76	1075.89
15	−1	+1	+1	+1	1362.09	1361.05
16	+1	+1	+1	+1	1052.86	1047.86

Similar results may be drawn by estimating the metamodel of Eq. (8) for *BIC* through statistical analysis. The OLS estimates are based on the results presented in Table 4 and are shown in Table 5. In Table 5, the second column refers to the estimated coefficients of the metamodel. The third column refers to the test statistic for the hypothesis *t*-test, in which the null hypothesis is $H_0 : \beta_{\cdot} = 0$ against alternative hypothesis $H_1 : \beta_{\cdot} \neq 0$, where $\beta_{\cdot}$ refers to a coefficient of the metamodel. For instance, β_{MS}, β_{PI}, β_{CI}, and β_{CA} stand for the coefficient of the metamodel related to MS, PI, CI, and CA, respectively. Furthermore, $\beta_{MS \cdot PI}$ refers to the coefficient related to the interaction $MS \cdot PI$; $\beta_{MS \cdot CI}$ refers to the interaction $MS \cdot CI$ and so on.

The statistical analysis of Table 5 reveals very close conclusions gathered by the main effect plots. From the table, one can see that the factor most important to *BIC* specification is the environmental variable MS, followed by PI. Looking at CI, statistical analysis shows a slightly, but statistically significant negative impact on information index *BIC*, improving its specification. As for environmental variable CA, it seems to not bear significantly on *BIC*, since CA's t-value is not significant in testing hypothesis $H_0 : \beta_{CA} = 0$. The results on CA suggest that it should be dropped from the (linear) model.

Interestingly, similar to what has been addressed also by Yu et al. (2009) for UK electricity utilities, the weather condition is significant on efficiency performance. Furthermore, one can see from Eq. (7) that, on average, the impact of weather condition in efficiency performance is small,

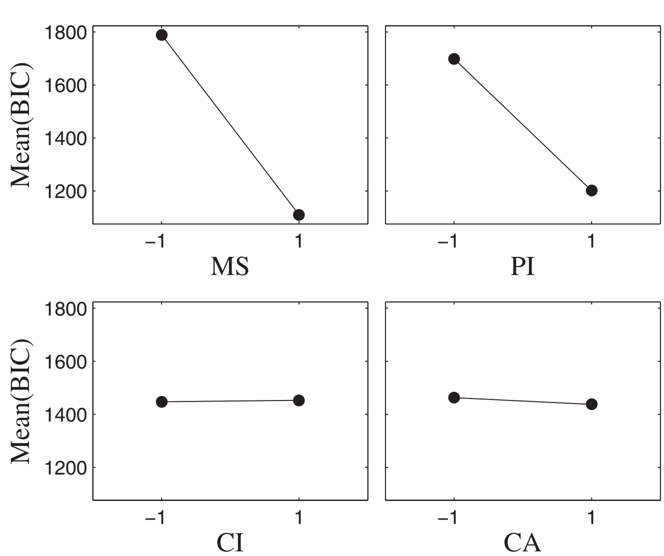

Fig. 1. Main effects for BIC.

Table 5
Parameter estimates for the metamodel.

	BIC	
	Coefficients	t-Values
(Intercept)	1449.6	614.77
β_{MS}	−340.6	−144.45
β_{PI}	−248.8	−105.49
β_{CI}	−13.9	−5.90
β_{CA}	1.7	0.73
$\beta_{MS \cdot PI}$	219.3	92.99
$\beta_{MS \cdot CI}$	−13.3	−5.63
$\beta_{MS \cdot CA}$	−13.7	−5.81
$\beta_{PI \cdot CI}$	5.8	2.47
$\beta_{PI \cdot CA}$	6.9	2.94
$\beta_{CI \cdot CA}$	−2.0	−0.83

suggesting that PI is internalized by DSOs in Brazilian case. As for CI, the Brazilian regulator advocates, in its technical report (ANEEL, 2011), that DSOs have a strong capacity to manage their non-technical losses, i.e., losses resulting from theft and measurement errors. The results on CI show a similar conclusion, since it has a small impact on *BIC* specification. Still, note that from all the environmental variables considered by Brazil's regulator, only CA does not improve *BIC* specification.

However, besides the main factor analysis discussed above, one may gather information about interactions between two environmental variables, i.e., when one environmental variable potentiates another. Such is the case of MS and CA, which seems to constitute an important interaction in the model. This interaction explains why CA appears to be significant in the linear model proposed by Brazil's regulator. Contrary to a linear impact on the efficiency score, CA reveals a nonlinear influence on efficiency performance, depending on the MS level. Furthermore, concerning efficiency performance, the interaction between MS and CI is also statistically significant. Importantly, as all these interactions present negative coefficients, they improve *BIC* specification. Finally, interactions are also observed between PI and CI and between Pi and CA. Nevertheless, these interactions worsen *BIC* specification, suggesting that it must not be considered in the econometric model.

Finally, based on the discussion above, one can thus estimate the second stage model considering the main effects (linear coefficients), except for CA, and interactions MS.CI and MS.CA:

$$\delta_i = -9.57 + 1.38 z_{MS} + 0.28 z_{PI} + 14.38 z_{CI} - 1.72 z_{MS} z_{CI} + 0.03 z_{MS} z_{CA} + \varepsilon. \tag{9}$$

The econometric model of Eq. (9) is significantly different from the approach used by Brazilian regulator given in Eq. (5). In Eq. (9), for instance, environmental variable CA nonlinearly affects the DSOs performance, depending on the MS level. Furthermore, besides the linear relationship between efficiency score, CI also interacts to MS level in affecting the DSO performance. Note also that, as previously discussed, the impact of weather condition in efficiency performance is small on average. The index criteria of the model of Eq. (9) is $BIC = 1001.4$, which is lower than one of Eq. (5), yielding a better suited model.

Fig. 2 depicts a radar plot comparing the second stage efficiency score of Brazil's regulator model and the one suggested by DOE methodology. The center of radar plot refers to the zero efficiency, whereas the fifth circle represents 100% efficiency. In mean, the difference between both models is about 2.1% annual permitted Opex. Although the difference between two models seems to be small, it may represent a difference of about \$ 205 million on annual permitted operational and maintenance costs between Brazil's regulator model and the one selected by DOE approach. Of these, about \$ 187.5 million would represent an annual deficit for electricity utilities, i.e., once the semi-parametric two-stage benchmarking model is adopted, the misspecification of the second stage may represent about \$ 187.5 million less than necessary income for distribution network operation and maintenance. Indeed,

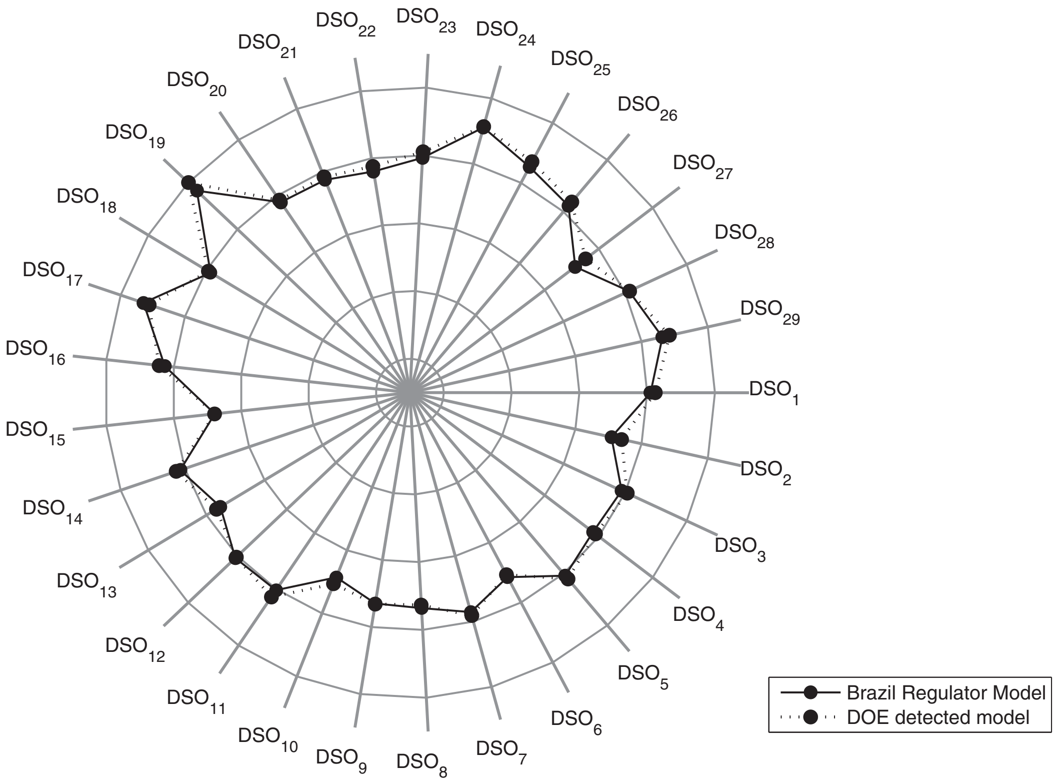

Fig. 2. Radar plot comparing both models.

for some DSOs, such difference in the models may represent up to 7.5% deficit in the allowed operational expenditures.

The results above suggest that Brazil's regulator made a fair choice on environmental variables, but the adopted model seems to be misspecified. Brazil's regulator defined a linear dependence among efficiency performance and environmental variables. The contrary is observed, however, when using DOE philosophy as a systematic simulation approach for environmental variables analysis. The results show that interactions between variables must be considered yielding a nonlinear model when analyzing environmental variables.

6. Conclusions

This paper has introduced a simulation approach based in design of experiments philosophy to analyze the second stage of semi-parametric two-stage approaches in the analyses of electricity utility companies' efficiency. Such analyses are commonly used by the regulator to set DSO's allowed revenue. In this paper, the two-stage DEA, which is composed of a traditional DEA model (first stage) and a regression analysis considering environmental variables (second stage), is used. To enable statistical inference in the censored regression, a bootstrap approach is considered.

In this context, the use of simulated design of experiments is proposed to evaluate the impact of environmental variables on decision-making units' efficiency, where a full factorial design is used. Using a metamodel for the efficiency score, this paper shows that, depending on contextual factor information, the criteria index may be affected differently. Thus, the regulator, in selecting environmental variables, should establish a clear objective.

To help with the efficiency analysis of environmental variables, Bayesian Information Criteria is used, a criteria widely used in selecting regression models. Such an index is a consistent information criteria and is based on the assumption that the data comes from a finite order process and has the probability of obtaining the true order of the model that, when the sample size increases, goes to one.

The results obtained shows that Brazil's regulator has fairly chosen environmental variables. However, the regression model selected may be misspecified. Instead of a linear dependence among efficiency performance and environmental variables, the simulated experiments (with the help of full factorial designs) reveal a nonlinear relationship, due to interactions among variables. Finally, it is important to note that the analysis developed in this paper was carried out with real data available on the regulator's website, and the results may be reproduced as the dataset used is available in the supplementary data of this paper.

Acknowledgments

The authors would like to thank CAPES, CNPq, and Fapemig, for financial support. The authors also thank anonymous reviewers for their valuable comments which enriched the paper discussion.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.eneco.2014.06.017.

References

Ajodhia, V.S., 2006. Regulating Beyond Price Integrated Price-quality Regulation for Electricity Distribution Networks(PHD thesis in the) Technische Universiteit Delft.
Ajodhia, V., Hakvoort, R., 2005. Economic regulation of quality in electricity distribution networks. Util. Policy 13 (3), 211–221 (Special Issue on Utility Service, Quality).
ANEEL, 2010. Technical Study Number 365/2010 (in Portuguese)Available at: http://www.aneel.gov.br/aplicacoes/audiencia/dspListaDetalhe.cfm?attAnoAud=2010&attIdeFasAud=432&id_area=13&attAnoFasAud=2011.
ANEEL, 2011. Technical Study Number 298/2011 (in Portuguese)Available at: http://www.aneel.gov.br/aplicacoes/audiencia/dspListaResultado.cfm?attAnoAud=2010&attIdeAud=462&attAnoFasAud=2011&id_area=13.
Balestrassi, P.P., Popova, E., Paiva, A.P., Lima, J.W.M., 2009. Design of experiments on neural network's training for nonlinear time series forecasting. Neurocomputing 72, 1160–1178.
Billinton, R., Allan, R., 1984. Reliability Evaluation of Power Systems. Plenum Press, London.

Brophy Haney, A., Pollitt, M., 2009. Efficiency analysis of energy networks: an international survey of regulators. Energy Policy 37 (12), 5814–5830.

Chan, K.K., Spedding, T.A., 2001. On-line optimization of quality in a manufacturing system. Int. J. Prod. Res. 39 (6), 1127–1145.

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

Chilingerian, J.A., Sherman, H.D., 2004. Health care applications: from hospitals to physicians from productive efficiency to quality frontiers. In: Cooper, W.W., Seiford, L.M., Zhu, J. (Eds.), Handbook on Data Envelopment Analysis. Kluwer Academic Publishers, Boston, pp. 265–298 (Chapter 10).

Coelho, J., Nassar, S.M., Gauche, E., Ricardo, V.W., Queiroz, H.L., de Lima, M., Lourence, M.C., 2003. Reliability diagnosis of distribution system under adverse weather conditions. Power Tech Conference Proceedings, 2003 IEEE Bologna, vol. 4, pp. 290–294.

Cooper, W.W., Seiford, L.M., Tone, K., 2007. Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software, 2nd edition. Springer.

Domijan Jr., A., Matavalam, R.K., Montenegro, A., Willcox, W.S., 2003. Analysis of rain, wind and temperature effects on power distribution outages. Proceedings of the IASTED International Conference, Power-Con-Special Theme: Blackout, pp. 44–48.

Evans, L., Guthrie, G., 2006. Incentive regulation of prices when costs are sunk. J. Regul. Econ. 29 (3), 239–264.

Farrell, M., 1957. The measurement of productive efficiency. J. R. Stat. Soc. A 120, 253–281.

Fisher, R.A., 1966. The Design of Experiments, 2nd ed. Hafner Publishing Company, New York (orig. pub. 1935).

Jamasb, T., Pollitt, M., 2001. Benchmarking and regulation: international electricity experience. Util. Policy 9 (3), 107–130.

Jamasb, T., Pollitt, M., 2003. International benchmarking and regulation: an application to European electricity distribution utilities. Energy Policy 31, 1609–1622.

Joskow, P., 2008. Incentive regulation and its application to electricity networks. Rev. Netw. Econ. 7 (4), 547–560.

Kleijnen, J.P.C., 2005. An overview of the design and analysis of simulation experiments for sensitivity analysis. Eur. J. Oper. Res. 164, 287–300.

Lowry, M.N., Getachew, L., 2009. Statistical benchmarking in utility regulation: role, standards and methods. Energy Policy 37, 1323–1330.

Montgomery, D.C., 2009. Design and Analysis of Experiments, 7th edition. John Wiley & Sons Inc.(656 pp.).

Odeck, J., 2009. Statistical precision of DEA and Malmquist indices: a bootstrap application to Norwegian grain producers. Omega 37 (5), 1007–1017.

Oliveira, F.A., Paiva, A.P., Lima, J.W.M., Balestrassi, P.P., Mendes, R.R.A., 2011. Portfolio optimization using mixture design of experiments: scheduling trades within electricity markets. Energy Econ. 33, 24–32.

Ray, S.C., 2004. Data Envelopment Analysis: Theory and Techniques for Economics and Operations Research. Cambridge University Press, Cambridge.

Ruggiero, J., 2004. Performance evaluation in education: modeling educational production. In: Cooper, W.W., Seiford, L.M., Zhu, J. (Eds.), Handbook on Data Envelopment Analysis. Kluwer Academic Publishers, Boston, pp. 265–298 (Chapter 10).

Sadjadi, S.J., Omrani, H., 2008. Data envelopment analysis with uncertain data: an application for Iranian electricity distribution companies. Energy Policy 36 (11), 4247–4254 (Transition towards Sustainable Energy Systems).

Shuttleworth, G., 2005. Benchmarking of electricity networks: practical problems with its use for regulation. Util. Policy 13 (4), 310–317.

Simar, L., Wilson, P.W., 2007. Estimation and inference in two-stage, semi-parametric models of production processes. J. Econ. 136 (1), 31–64.

Schwarz, G., 1978. Estimating the dimension of a model. Ann. Stat. 6 (2), 461–464.

Steele Santos, P.E., Leme, R.C., Galvão, L., 2012. On the electrical two-part tariffs — the Brazilian perspective. Energy Policy 40, 123–130.

Sun, Y., Li, X., 2013. Optimizing surgery start times for operating room via simulation. Int. J. Healthc. Technol. Manag. 14 (1/2).

Thanassoulis, E., Portela, M.C.S., Despic, O., 2008. Data envelopment analysis: the mathematical programming approach to efficiency analysis. In: Fried, H.O., Lovell, C.A.K., Schmidt, S.S. (Eds.), The Measurement of Productive Efficiency and Productivity Growth. Oxford University Press.

Tsionas, E.G., Papadakis, E.N., 2010. A Bayesian approach to statistical inference in stochastic DEA. Omega 38 (5), 309–314 (Empirical Research in the EU Banking Sector and the Financial Crisis).

Wang, P., Billington, R., 2002. Reliability cost/worth assessment of distribution systems incorporating time-varying weather conditions and restoration resources. IEEE Trans. Power Deliv. 17 (1), 260–265.

Yu, W., Jamasb, T., Pollitt, M., 2009. Does weather explain cost and quality performance? An analysis of UK electricity distribution companies. Energy Policy 37 (11), 4177–4188.