UCLouvain

LOUVAIN
School of Management

Louvain School of Management

Analysis of the benchmarking models used for the regulation of DSO's in the European Union

Implementation of the StoNED model in R

Author : Augustin Van Bavel
Supervisor : Per Joakim Agrell
Academic year 2019-2020.

I would first like to thank my thesis advisor Professor Per Joakim Agrell of the Louvain School of Management at Université Catholique de Louvain.

I must express my very profound gratitude to my parents and to my partner for providing me with unfailing support and continuous encouragement.

Abstract

This master thesis analyses three different benchmarking methods used in different countries for the regulation of electricity distribution system operators (DSO's) designed to promote efficient energy network markets. The three methods we will focus on are (i) the data envelopment analysis (DEA), (ii) stochastic frontier analysis (SFA) and (iii) stochastic non-parametric envelopment of data (StoNED). The goal of this master thesis will be to implement the routines in R allowing the StoNED model to be computed and compare it to the two other models. R is a free software environment for statistical computing.

First the three methods will be theoretically presented along with the general ideas behind benchmarking techniques. The three models will then be tested using the same data from the Swedish DSO's from 2001 to 2006. The results obtained will be discussed, and compared against each other. The results obtained for the StoNED model will be the focus of this study as the two other models are already well known and have a dedicated package, that will be used to analyse the DEA and SFA models.

The resulting scores that were computed for the StoNED model are for most of the cases correlated with the scores computed using the two other models. This seems to indicate that the implementation of the StoNED model in R is reliable. If the StoNED model is proved accurate, it will be able to combine the main advantages of both DEA and SFA, namely a flexible functional form and the consideration of noise in the data. Additionally, some limitations that were encountered when implementing the model in R will be discussed.

Contents

Introduction

In its consultation paper of 2017 [2], the Council of European Energy Regulators (CEER) stated that: "The regulation should be designed so that it maintains incentives for cost-efficiency, quality of service and supply, and guaranteed financial viability for efficient DSOs."

It is crucial to keep a well-designed incentive scheme in order to have an efficient energy network market. The regulation methods need promote cost efficiency, while ensuring the financial viability of the firms it regulates.

This master thesis will analyse three different benchmarking methods. The first two are well-known and are currently used in Europe by regulators. The third one is more advanced, as it combines properties of the first two methods to try to tackle the main drawbacks of these techniques.

Could a semi-parametric approach like StoNED be used by regulators in place of DEA and SFA for the benchmarking of DSO's? This master thesis will compare the results obtained for the three methods when applied on the network data of the Swedish DSO's between 2000 and 2006.

The three benchmarking presented in this thesis are going to be tested on a data set consisting of previous data from the Swedish DSO's. The different algorithms used to compute the performance scores of each firm will be compared. The validity of the methods employed is crucial as a small difference in an efficiency score could lead to a huge loss of income for the DSO's.

Chapter 1

Context

The market of electricity distribution is a natural monopoly. Traditional competition cannot be applied due to high barriers to entry, such as excessively large investment costs, tiny marginal costs, and often the impossibility to run multiple wires through a same area. These particularities makes that the consumer has to supply its electricity from a specific distributor (hereafter referred to as DSO), depending on its location, and cannot choose between different distributors. The lack of competitive pressure could lead to voluntary inefficiencies from DSO's. The actions they could carry out could be a reduction of the costs, accompanied with a lower quality of service.

To introduce a first layer of competition, the European Commission [32] introduced the "unbundling" principle in order to separate the different stages of the supply chain, thus creating competition on the vertical level. This principle forced the energy supplier to be different from the energy distributor, which is the reason now the energy bill is separated into two distinct parts, the cost of energy consumed and the cost of delivering the energy.

European governments have also taken measures to counteract this phenomenon on the horizontal level, to introduce competition between distributors operating on the same level. Because the different distributors each operate in different locations, they benefit from a natural monopoly. This type of monopoly being legal for the reasons mentioned above, regulators have been put in place in order to regulate the players in this market.

Different countries have set up different regulatory environments. The various systems which have been installed range between light and heavy handed regulations. A presentation of the different models put in place can be found in Agrell, Teusch (2015) [6]. The different regulatory mechanisms will be presented here.

1.1 Regulatory Mechanism

A regulatory mechanism can be defined by four parameters of the model: (1) the incentive power, (2) the regulatory updating, that is the length of the regulatory period, (3) the revenue function, which defines the revenue model of the firm, accounts for inflation and cost elements that might not be included in the cost function and finally, (4) the cost model. The cost model can be defined as in Agrell, Teusch (2015) [6] : "All possible engineering/economic, deterministic/stochastic, parametric/non-parametric, partial/total techniques to estimate the cost or cost development of defined operation given some set of relevant data". The techniques to estimate the cost model that will be investigated here are the Data Envelopment Analysis (DEA), the Stochastic Frontier Analysis (SFA) and the Stochastic Nonparametric Envelopment of Data (StoNED).

However, it is the revenue function that will determine the tariffs a DSO will be able to charge. The main regimes of revenue function used in Europe are the price cap and revenue cap.[1]

1.2 Regulatory Regimes

There exists multiple techniques for the regulator to control the operators. They range from light-handed to heavy-handed, and have different incentive power. A list of the different techniques used by regulator can be found in the annex 1.1. It was put together in a consultation paper by the council of European energy regulators.

[1] Agrell, Teusch (2015) [6]

Cost-plus

One of the most basic one is the cost plus regime. With this model, all the costs observed *ex post* are reimbursed, and the capital invested is also taken into account. The incentive power is very low as the operator does not face any risk. This regime has a considerable burden as there exists a large information assymetry between the regulator and the operator. It is entirely the role of the first to ensure that the second does not inflate costs inappropriately.

Price/revenue cap

Another typical regime is the price cap. The price cap regulation restricts the price a regulator can charge. The price cannot be increased more than a certain amount and the cap is reviewed for each regulatory period. The impact of the price increase is defined by two factors: the inflation (I) and a "productivity offset"[17] (X). The productivity offset will be the main focus of this thesis. It is calculated using the different models that will be presented in the next sections. The main burden of this method is that once a cap has been set, it is applied for the whole regulatory period (depends on the country, usually four to five years). Due to the ratchet effect (the impossibility to come back on these decisions), the regulated firm and the regulator have to properly anticipate the next period. If the cap is too high, the DSO will make too much profit, if it is too low, the firm will face financial difficulties.

A difference can be distinguished between the *pure* price cap and the *hybrid* price cap. In the first, the regulator does not have any access on the actual profit made by the operator. The only inputs are *ex ante* information. With the hybrid price cap or revenue cap regime, the regulator takes into account some *ex post* information to adjust the calculated price. Jamison (2005) [17] states the three main advantages of using such a regime in regulation: they are "(1) providing companies with incentives to improve efficiency, (2) dampening the effects of cost information asymmetries between companies and regulators, and (3) decreasing the incentives to over-invest in capital and cross-subsidize relative to rate of

Country	Regime	Method DSO	Method TSO
Austria	Revenue cap	DEA-MOLS(nat)	DEA(int)*
Belgium[a]	Revenue cap	DEA(nat)*	DEA(int)
Denmark	Revenue cap	MOLS(nat)	DEA(int)
Estonia	Revenue cap	MOLS(nat)	DEA(int)*
Finland	Revenue cap	StoNED(nat)	DEA(int)
France	Cost recovery	Ad hoc	DEA(int)*
Germany	Revenue cap	DEA-SFA(nat) best-of	DEA(int)
Great Britain	Revenue cap	MOLS(nat)	DEA(int)*
Greece	Cost recovery	–	DEA(int)*
Hungary	Price cap	Ad hoc	Ad hoc
Iceland	Revenue cap	Ad hoc DEA(int)*	DEA(int)
Ireland	Price cap	Ad hoc	Ad hoc
Italy	Revenue cap (opex)	Ad hoc	DEA(int)*
Lithuania	Price cap	Ad hoc	DEA(int)*
Luxemburg	Cost recovery	Ad hoc	DEA(int)*
Netherlands	Yardstick	MOLS(nat)	DEA(int)
Norway	Yardstick	DEA(nat)	DEA(int)
Portugal	Revenue cap	SFA(nat)	DEA(int)
Spain	Revenue cap	Engineering	DEA(int)*
Slovenia	Price cap	DEA(nat)	–
Sweden	Rate of return	Ad hoc DEA(nat)*	DEA(int)*
Switzerland	Cost recovery[b]	Ad hoc DEA(nat)*	–

Figure 1.1: Regulatory Regimes

Source: Agrell and Niknazar, 2013

return regulation."

Most regulators use the hybrid price cap/ revenue cap, as it is somewhat in the middle of a pure price cap and a cost-plus regime.

Yardstick regulation

When the result of a benchmarking study is used to allocate more or less resources to a firm, it is a case of yardstick regulation.

This type of regime is heavy-handed, as it requires at first a considerable amount of information to properly benchmark the regulated entities. However, it provides these entities with the most incentive power compared to the previous models. While only a few European countries operates a true yardstick regulatory regime, a significant amount of regulators use parts of yardstick regulations in their approaches, being in a price cap regime to identify the "productivity offset", or simply as an informational tool. 1.1 shows the different regimes adopted by the European regulators.

1.3 DSO's in Belgium

In Belgium, DSOs are ORES, Tecteo (Resa), Régie de Wavre, AIESH and AIEG in Wallonia, Sibelga in the Brussels-Capital Region and Eandis and Infrax in Flanders. The responsibility of regulation has been passed to the regions since the sixth state reform in 2014. It was previously operated by the national regulator, the CREG. The regulators are now the CWaPE in Wallonia, the VREG in Flanders and Brugel in Brussel.

This split of the competences has made the implementation of an efficient yardstick regime even more difficult, as the number of regulated entities per regulator has decreased. Agrell and Teusch (2015) [6] make the case for an harmonization between regions/countries, allowing the regulators to perform benchmark studies on national/international DSO's.

Chapter 2

Literature Review

This chapter will consist of a description and analysis of the three models that will be discussed in this master thesis. An introduction on the use of benchmarking methodology will first be discussed. Each model will then be presented, from its origins to their current state. Their assumptions will be reviewed, and their current uses in the electricity distribution sector will be discussed.

The models that will be discussed here have received varying levels of discussion in the scientific literature. DEA has been heavily reviewed, so the scope of this master thesis will lean more toward the SFA and StoNED models. Nonetheless, the DEA model will be thoroughly analyzed in the literature review, as it is currently the most used method in the regulation of DSO's in Europe.

In addition, the different methods of regulation of the electricity distribution sector will be reviewed and compared.

2.1 Benchmarking Theory

This section will provide an overview of some of the basic concepts of benchmarking, such as simple efficiency measures, technology frontier and more. Most of these core

subjects from the Benchmarking theory are summarized from the Bogetoft and Otto (2011) [10].

2.1.1 Efficiency

In this section the efficiency of a unit will be defined, and the key measures of this efficiency will be presented. The word unit can refer to any separate entity that we want to benchmark, it can equally represent a whole firm or a division of a company that will be compared to other divisions. Because this master thesis deal with distribution operators, who are separate entity, the terms "units" and "firms" will be used interchangeably.

The efficiency of a unit can be defined as the ratio of its outputs over its inputs. Calculating the efficiency of a firm producing one type of output using only one type of input is very simple, it can be expressed as a simple KPI. However, it quickly become more complex when multiple inputs and outputs are at play. This method of describing the efficiency simply as a ratio composes several limitations, as exposed in Bogetof and Otto (2011) [10].

- Constant Return To Scale : When using simple ratios, we have to assume constant return to scale. In reality, an increase in the scale of the production can result in either some economies of scale or inefficiencies, that makes the return to scale increasing or decreasing
- Partial evaluations : When different inputs are used to produce one output, it becomes possible to compute different KPI's. Thus the most efficient unit could differ depending on the KPI used.

2.1.2 Technology Set

In the benchmarking theory, the technology set is the set of all input/output combinations. If we do not have prior information about the technology set, we can nonetheless

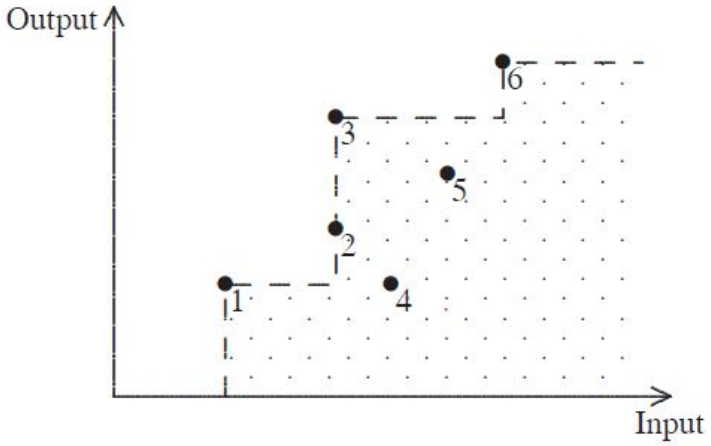

Figure 2.1: Free Disposability

suppose that it contains all the current observations. From there, we can estimate a frontier that will be the limit of the technology set. This frontier can vary depending on the assumptions used in its computation.

Free Disposability The first assumption, free disposability, implies that for each observed input/output combination, any combination using more inputs or producing less outputs is also feasible. Using only this assumption, the implied technology set can be visualized as in Figure 2.1.

Convexity This second assumption implies that if two units have different input/output combination, then a mix of these two units is also feasible. This is called a convex combination. If we assume free disposability and convexity, the approximated technology set can be represented as in figure 2.2

The efficiency can be defined using this concept, as the most efficient units will be present on the frontier of the technology set. By computing the distance between each unit and this frontier, we can have an overview of the efficiency.

The frontier of the technology set can also be represented by an isoquant. On a graph where each axis represent an input (or an output), the isoquant is the line of the best

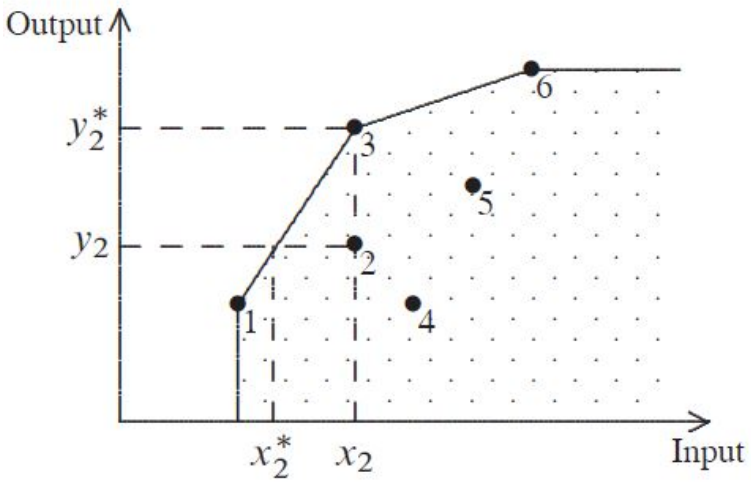

Figure 2.2: Free Disposability And Convexity

possible product mix used for a fixed output (or best possible output for a fixed input).

2.1.3 Farrell Efficiency

In his paper of 1957, "The measurements of productive efficiency", M. J. Farrell [15] presents a simple way to evaluate a unit's efficiency. He defines the technical efficiency as the factor that multiplies a combination of inputs or outputs so that the result is on the efficient production function. This measure of the efficiency is used as a core concept in the modern benchmarking theory.

By taking the Farrell efficiency as defined by Bogetoft and Otto [10], we can mathematically define the Farrell measures as follows:

- Input Farrell Efficiency :

$$E = min\{E > 0 | (Ex, y) \in T\} \tag{2.1}$$

- Output Farrell Efficiency :

$$F = max\{F > 0 | (x, Fy) \in T\} \tag{2.2}$$

$x \in \mathbb{R}_+$ representing the set of inputs, $y \in \mathbb{R}_+$ the set of outputs, and T the technology set frontier, or efficient production function.

An alternative to the Farrell efficiency is the Shepard efficiency. This measure is the the inverse as the Farrell one and thus does not bring any more information. It is sometimes seen in the literature or in some computations. As it will be used later in the analysis, the equations will be presented here:

$$\begin{aligned} D_i(x,y) &= max\{D > 0 | (\frac{x}{D}, y) \in T\} = \frac{1}{E(x,y)} \\ D_o(x,y) &= max\{D > 0 | (x, \frac{y}{D}) \in T\} = \frac{1}{F(x,y)} \end{aligned} \tag{2.3}$$

Another method is the directional efficiency measure. Unlike the Farrell efficiency, it doesn't multiply all the inputs/outputs with the same parameters. This allows a model to compute the efficiency by moving toward the technology set frontier in different directions, thus computing different changes in inputs/outputs. Hence it increases the learning opportunities.

2.1.4 Technical Efficiency, Allocative Efficiency And Cost Efficiency

There exists different forms of efficiency. The ones defined above are defined as the technical efficiency, as described by Farrell. The problem with technical efficiency is that it takes as information only the quantity of a certain input, without taking into account the price.

To circumvent this problem, the allocative efficiency and cost efficiency can be used. It makes use of a new parameter, a vector $w \in \mathbb{R}_+$ representing the prices of each input or output. The new efficiency measure doesn't rely thus solely on the x or y but on wx or wy.

The technical efficiency can be redefined as the efficiency of the unit compared to the best possible scenario using the same ratio of inputs/outputs, $\widetilde{x}$, the cost efficiency is the efficiency associated with the best possible allocation of resources, x^*. The Cost Efficiency can be defined by the following ration:

$$CE = \frac{wx^*}{wx} \tag{2.4}$$

where x^* is the input mix where the isoquant is tangent with the isocost curve, hence the cheapest way to produce the fixed output.

The allocative efficiency is the comparison between the best possible input mix while keeping the same input ration and the best and cheapest input mix for a fixed output. It can be mathematically defined as:

$$AE = \frac{wx^*}{w\widetilde{x}} \tag{2.5}$$

The cost efficiency is equal to the multiplication of the allocative efficiency and the technical efficiency.

2.2 Data Envelopment Analysis

In this section, the Data Envelopment Analysis (DEA) method will be described. It is the most used method to benchmark the electricity distribution sector, and is used in numerous other applications. The concepts, pros and cons of this method will be exposed bellow, but as this method has received extensive research, it will not be discussed in depth. Let's just state that DEA models have evolved, and now a lot of different ones exist besides the first one introduced by Charnes, Cooper and Rhodes in 1978 [12]. The one that will be presented here follows the definition of Bogetoft [10].

The advantage of DEA, compared to a standard computation of the efficiency described above is that it compares units that are comparable. It will always compare a firm to the one that is the most closely related on the technology set frontier.

The DEA method is a deterministic, non-parametric method. We will discuss the pros and cons of this method in the subsections below.

2.2.1 Model Definition

Model

This section will describe the DEA model for the input efficiency. The orientation of the model defines the type of efficiency which is studied. It can be either input or output oriented. The method to calculate the output efficiency will be quickly explained, as it is very similar to the one described bellow.

Let define $K \in \mathbb{N}_+$ as the number of firms. Each firm uses a vector of inputs x to produce a vector of outputs y. The *input* efficiency measure E_k represent the benchmarking score associated to a firm k. This efficiency measure can be computed as follow:

$$\begin{aligned} \min_{E,\lambda_1,\ldots,\lambda_k} \quad & E \\ \text{s.t.} \quad Ex^0 \quad &\geq \quad \textstyle\sum_{k=1}^{K} \lambda_k x_k; \\ y^0 \quad &\leq \quad \textstyle\sum_{k=1}^{K} \lambda_k y_k; \\ \lambda \quad &\in \quad \Lambda_K(\gamma); \end{aligned} \tag{2.6}$$

The parameter λ represents the weights of the input and output of each firm and determines to which other units the one we are investigating will be compared. The last constraint determines the structure of the return-to-scales. The different assumptions regarding the return to scales will be defined later.

If one wants to compute the output efficiency, the efficiency term E needs to divide the left-hand side of the second constraint.

Technology Set

The DEA model relies on the fact that the functional form of technology set (T) is not an assumed *a priori*, and is computed using endogenous weights. The approximation of the set (T^*) is done using the *minimal extrapolation principle*. This principle states that T^* must be as small as possible and contain all the observations. Nonetheless, some assumptions still need to be made in order to approximate the technology set. In order to define the most accurate set possible, four assumptions are expressed:

- Free disposability: As stated in a previous section, the approximation of the technology set assumes the free disposability of the inputs. We can achieve the same output with more inputs or a smaller output with the same input. Mathematically: $(x, y) \in T, x' \geq x, y' \leq y \Rightarrow (x', y') \in T$
- Convexity: Any production plan that is a composition of two existing production plan is feasible: $(x, y) \in T, (x', y') \in T, \alpha \in [0, 1] \Rightarrow \alpha(x, y) + (1 - \alpha)(x', y') \in T$
- γ-returns to scale: The production can be scaled but needs to respect the assumption made about the returns to scale: $(x, y) \in T, k \in \Gamma(\gamma) \Rightarrow k \cdot (x, y) \in T$
- Additivity: Two feasible production plans that are merged produce a feasible production plan: $(x, y) \in T, (x', y') \in T \Rightarrow (x + x', y + y') \in T$

The third assumption, γ-return to scale, states the way a production plan can evolve when it is scaled. The exist different assumptions that have been made through the DEA literature, we will present here six that are presented in Agrell (2015) [3].

Return to scale	Constraint
constant	$\{\lambda_k \geq 0 : k \in K\}$
non-increasing	$\{\sum_{k \in K} \lambda_k \leq 1, \lambda_k \geq 0 : k \in K\}$
non-decreasing	$\{\sum_{k \in K} \lambda_k \geq 1, \lambda_k \geq 0 : k \in K\}$
variable	$\{\sum_{k \in K} \lambda_k = 1, \lambda_k \geq 0 : k \in K\}$
free replicability	$\{\lambda_k = N_+ : k \in K\}$
free disposability	$\{\sum_{k \in K} \lambda_k = 1, \lambda_k = \{0, 1\} : k \in K\}$

(2.7)

source: [**?**]

The different assumptions impacts the size and shape of the technology set. The argument to favour one assumption rather than an other is important as it can lead to varying efficiency scores when comparing the same units. The units will prefer the return-to-scale structures that minimize the size of the technology set, as it will place them closer to the frontier. On the other side, the assumption that leads to the largest technology set, the constant return to scale, will give the units the worst score of all the assumptions.

Peer Units

When applying the DEA method, a firm is not compared to all the other firms, but rather related to a few peer units which are on the frontier and whose combination of inputs and outputs matches the most the one of the analysed firm. This observed firm is in fact compared to a reference firm, a fictive unit computed by combining the weight and production plan of each other unit. Based on the model ADD NUMBER, the production plan of firm j can be expressed as

$$(x_j, y_j) = (\sum_{k=1}^{K} \lambda_k y_k, \sum_{k=1}^{K} \lambda_k x_k) \tag{2.8}$$

λ_k being the weight associated to each firm k, the peer units of a firm j can be defined as the firms whose weights are greater than 0, more formally, the peer units of firm j are $\{k \in \{1, ..., K\} | \lambda_k > 0\}$.

The DEA method thus gives us a clear and easy method to find the true comparable competitors of a specific firm, and gives us the relative efficiency of the firm.

Scale Efficiency

The assumptions made *a priori* on the structure of the return to scales also have an impact on the way a firm will evolve if we rescale the cost attributes. If proportion between inputs and outputs are the same when the firm is scaled, the model follows a CRS assumption. In fact, the return to scale can be guessed when the CRS, IRS or DRS assumptions are made. But in the case of the VRS, one can ask how the efficiency of a firm will evolve if we scale the production up. To do so, it is useful to look at the average output per input. One can see that most often, when the VRS is assumed, the ratio is in order increasing, constant then decreasing. The optimal emplacement for a firm is thus to be on the line where the average output per input is constant. This point/area of a graph is called the *most productive scale size* (MPSS).

A firm that has a efficiency of 1 under the CRS assumption model, is in the MPSS area. The concept of *scale efficiency* shows how far a firm is from the MPSS, it takes the ratio of the efficiency calculated using the CRS assumption on the efficiency found using VRS model. As the CRS will always shows a firm under its worst day, the scale efficiency will never be greater than 1.

Flexibility

The DEA model is deterministic and non-parametric. The fact that it is non-parametric makes it more flexible, as the functional form of the frontier of T is not an assumption made *a priori*. The model relies on the data to generate T^*. As it gives more importance to the observations, it also make the model much weaker against outliers. Furthermore, the computed efficiency could be affected by environmental realities. In the case of the electricity network sector, one firm could for example work only with underground cables, and another only with overhead cables. The handling of the two type of networks being

different, it could thus lead to a biased efficiency measure if a strong return to scale assumption is made (CRS for example).

2.3 Stochastic Frontier Analysis

Besides DEA, the stochastic frontier analysis (SFA) is another approach that has been proven popular for estimating the DSO's performance. While not being used as widely as the previous method, it presents some features that allows to counteract some limitations of the DEA.

The SFA model is a parametric one. This means that the functional form of the distribution function is determined *a priori*. The unkonwn parameters of this functional form are represented by a vector β. This vector is unknown, but it can be used to determine the production function $f(x;\beta)$.

2.3.1 Noise and inefficiency

One of the core differences between DEA and SFA is that the later takes into consideration the noise in the data. During the data generation process, some inaccuracies or errors could lead to biased data that do not reflect the reality of the situation. With the DEA model, the noise in the data is not taken into account. The technology set frontier will thus embrace all the observations and what should normally be considered as noise is considered as inefficiencies. The advantage of SFA is that both the noise and the inefficiency are accounted for.

The inefficiency u can only influence the output to be smaller that what it would be in the production function without the inefficiency. Whereas the noise v can influence in both direction the output. The production function thus takes the following formulation: $y = f(x;\beta) + v - u$.

2.3.2 Functional Form

As stated above, the functional form of the production function needs to be estimated. In order to determine the firm specific efficiencies, one need first to determine the production function, the part of it associated to random errors and the part associated to inefficiencies.

The production function of the base model for SFA has been defined above, in are formal manner, for K units, the production function is:

$$\begin{aligned} y^k = \quad & f(x_k;\beta) + v_k - u_k, \\ & v_k \sim N(0,\sigma_v^2), u_k \sim N_+(0,\sigma_u^2), \quad k = 1,...,K. \end{aligned} \tag{2.9}$$

The model can be divided in two part: one deterministic one, $f(x^k;\beta)$, and one stochastic one that represents the variation from the production frontier, $v^k - u^k$. The deterministic part can take various forms, but one standard production function that is used as an example in both Bogetoft (ADD DATE) and Coelli and al (1998) ADD FOOTNOTES is the Cobb-Douglas production function. The linear form of the Cobb-Douglas can be expressed as:

$$\ln y_k = \beta_0 + \beta_1 \ln x_k \qquad \forall k = 1,...,K \tag{2.10}$$

The stochastic part, $v_k - u_k$ can be grouped in one variable named $\epsilon_k = v_k - u_k$ representing the error term, the variation of the unit against the production function. The error terms are subject to some assumptions:[1]

- Independence: v_k and u_k are distributed independently from each other, and are not correlated to the input vector x_k.
- Zero mean: $E(v_k) = 0$,
- Homoskedasticity: $E(v_k^2) = \sigma_v^2$ and $E(u_k^2) = \text{constant}$,
- Uncorrelated: $E(v_i v_j) = 0 \quad \forall i \neq j$ and $E(u_i u_j) = 0 \quad \forall i \neq j$

[1] coelli

These assumptions concerning the error terms where first made by Aigner, Lovell and Schmidt. They are justified by the fact that firms faces an input quantity or measurement error, which is symmetric, and a error linked to their inability to follow the best practices, which is one sided.

It is worth noting that we can get insights from the variances of the errors. Indeed, if σ_u is larger than σ_v more of the deviation from the frontier is explained by the inefficiency of the units observed. On the other hand, if the relation is opposite, the deviation is mainly explained by the random noise.

Using Cobb-Douglas, the final production function takes the following form:

$$y_k = \exp(\beta_0 + \beta_1 \ln x_k) \times \exp(v_k) \times \exp(-u_k) \qquad \forall k = 1, ..., K \tag{2.11}$$

This equation gives the error terms a multiplicative impact. It is convenient as the Farrell efficiency can be more easily computed. The output Farrell efficiency defined earlier can be defined as the maximal output possible/actual input. Using the definition of the production function ADD NUMBER, the Farrell efficiency can be expressed as

$$F(x_k, y_k) = \frac{f(x_k; \beta) \exp(u_k)}{f(x_k; \beta)} = \exp(-u_k) \tag{2.12}$$

To correctly estimate the production frontier function, the maximum likelihood estimation method is applied. This method will find the values for the parameters that correspond the most to the observations, the values for the parameters β will thus be the values that maximize the log-likelihood function. used the parameters $\sigma^2 = \sigma_v^2 + \sigma_u^2$ and $\lambda^2 = \frac{\sigma_u^2}{\sigma_v^2}$ to find a log likelihood function for the half-normal distribution associated with the error term. Let's note again that solely based on the values of these parameters, one can already make some assumptions. Indeed, if λ is greater than one, the variation of the units from the production function will tend to be explained more by the inability to follow the best-practices. On the opposite side, if lambda is smaller than 1, the variation will be more related to the random noise in the data. It is often easier to work with log likelihood function, the one which will be used here is formed by Aigner, Lovell and

Schmidt (1977) [7] and expressed as

$$\ln L(y|\beta,\sigma,\lambda) = -\frac{K}{2}\ln(\frac{\pi\sigma^2}{2}) + \sum_{k=1}^{K}\ln\Phi(-\frac{\epsilon_k\lambda}{\sigma}) - \frac{1}{2\sigma^2}\sum_{k=1}^{K}\epsilon_k^2 \tag{2.13}$$

where y is the log vector of the outputs and $\Phi(x)$ is the cumulative distribution function of the standard normal distribution. Different methods exists to solve this function, the method that will be used here is the same as in Bogetoft and Otto (2011), as the same package will be used in R for both SFA and DEA. The method permits to find the parameters β as well as

Among the parameters estimated through the log-likelihood estimation, the lambda parameter can be easily interpret to find the proportion of deviation from the production function that is due to the inefficiency factor. This proportion is found as

$$\frac{\lambda^2}{\lambda^2+1} \tag{2.14}$$

As the deviation is expressed by two factor, the rest of the deviation is explained by the random noise v, which is computed as $1-\frac{\lambda^2}{\lambda^2+1}$.

2.3.3 Determining the Efficiency

For now, only the proportion of the variation from the production function associated to inefficiencies and noise has been evaluated. Now that the functional form has been defined, the firm-specific efficiencies can be estimated.

$$\hat{\epsilon_k} = \ln y_k - \ln f(x_k;\hat{\beta}), \qquad \forall k = 1,...,K \tag{2.15}$$

The error term can be derived from the functional form once the parameters have been estimated, nevertheless the part of the error related to the inefficiency still cannot be derived. In order to obtain the value of u_k, the conditional estimation of u_k given ϵ_k will be used. However, the conditional density $\phi(u_k|\epsilon_k)$ is needed for the computation of the

conditional estimation. Fortunately, this density can be retrieved. Knowing that v_k and u_k are independent, the density of $\phi(u_k, v_k)$ is equal to $\phi(u_k)\phi(v_k)$. The term v_k can be replaced by $\epsilon_k + u_k$, the joint density can be expressed as $\phi(\epsilon_k + u_k, u_k) = \phi(\epsilon_k + u_k)\phi(u_k)$ and using Bayes theorem, the conditional density is written as

$$\phi(u_k|\epsilon_k) = \frac{\phi(\epsilon_k + u_k)\phi(u_k)}{\phi\epsilon_k} \tag{2.16}$$

To find the correct estimation of the technical efficiency, 2 auxiliary variables are used, they represent respectively the estimations of the mean and variance of the half normal distribution of the inefficiency term u_k:

$$\begin{aligned} \mu_* &= -\epsilon\frac{\lambda^2}{1+\lambda^2} \\ \sigma_* &= \frac{\lambda}{(1+\lambda^2)}\sigma \end{aligned} \tag{2.17}$$

Where ϵ is the vector of length K of the error terms for the K units. The conditional expectation of u_k given ϵ_k is written as

$$\mathrm{EV}(u|\epsilon) = \mu_* + \sigma_*\frac{\phi(\mu_*/\sigma_*)}{\Phi(\mu_*/\sigma_*)} \tag{2.18}$$

Now that the random variable u_k has been estimated, the efficiencies can be calculated using formula 2.18 However, better estimators exists. The estimator in equation 2.19 minimize the mean square error.

$$TE = \mathrm{EV}(\exp -u|\epsilon) = \frac{\Phi(\mu_*/\sigma_* - \sigma_*}{\Phi(\mu_*/\sigma_*)}\exp\left(\frac{1}{2}\sigma_*^2 - \mu_*\right) \tag{2.19}$$

2.3.4 Error distribution

In the section above, we assumed that the error term has a half-normal distribution. While this assumption is the one introduced by Aigner, Lovell and Schmidt (1977) [7], it is by no means the only one that can be considered. Coelli et al. (2005) [14]summarize the different assumptions that can be made, these assumptions will be presented in this section but

the application in this master thesis will stick with the half-normal distribution.

The truncated normal ($u_k \sim N^+(\mu, \sigma_u^2)$) is the closest to the half-normality assumption. The only difference is that instead of having a mean of 0, this distribution has mean μ. The inefficiency term still can only take a positive value. This assumption is due to Stevenson (1980) [31]. With the exponential assumption ($u_k \sim G(\lambda, 0)$), the inefficiency term follows the exponential distribution, with mean λ. The gamma model ($u_k \sim G(\lambda, m)$) makes u_k follow the gamma distribution, with a mean λ and m degrees of freedom.

2.3.5 Bayesian Stochastic Frontier Analysis

While not as widely used as the DEA method, the SFA is used by some national regulators to evaluate the best practices in their countries. One can find applications of this method in Germany, Portugal and Poland [2].

The polish application is particularly interesting as it uses bayesian model selection (BMS) and inference pooling techniques to identify the best explanatory variables. This method can prove useful when the model faces many uncertainties on the variables that should be included. This method is particularly appealing in the case of DSO regulation, following Makiela and Osiewalski (2018) [29] analysis:

- It allows to compare different models between them. The best among the competing models can then be chosen. The choice of variables to be included in the different models can have a large impact on the results.
- The bayesian inference method allows to reach exact results for small samples. This can prove very useful for some countries, like Belgium, which do not present a large number of DSO's.

The BMS method allows for a fair and transparent variable selection process, which can be critical when there are some conflicting interests and the regulator needs to be as

[2] [5] and [29]

transparent as possible.

The SFA model used by Makiela and Osiewalski is of the exponential type. The bayesian model is written as:

$$p(y, \beta, \sigma_v^{-2}, u, \phi|X) = p(\beta, \sigma_v^{-2}, \phi)\Pi_{i=1}^{n}\left(p(u_i|\phi)\Pi_{t=1}^{T} f_n(y_{it}|x_{it}\beta + u_i, \sigma_v^2)\right) \tag{2.20}$$

where ϕ is a parameter of u_k.

The different models have then been estimated using Markov Chain Monte Carlo Class Algorithm. Once the model have been defined, the authors have been able to compare them and pool inference. The sampled models are defined as

$$M_j : p_j(y|\lambda, \eta_j), \quad j = 1, ..., m, \tag{2.21}$$

The vector of unobserved quantities is defined as θ and is composed of two terms: λ represents the parameters common to all models and η which defines quantities specific to model M_j. The m bayesian models are expressed as $p_j(y, \theta_j) = p_j(y|\theta_j)p_j(\theta_j)$ and the probability for each model is defined by the bayes formula

$$p(M_j|y) = \frac{p(M_j)p(y|M_j)}{\sum_{i=1}^{m} p(M_i)p(y|M_i)} \tag{2.22}$$

This type of model was difficult to compute due to the inability of the model to find marginal data density (MDD) for latent variables. However, the authors point out that a new develoment made by Pajor (2016) allows them to find an unbiased estimator of MDD. In the bayes formula, the MDD is expressed as $p(y|M_j)$. One can thus notice that the probability of the model depends (i) on the prior probability of M_j and on the MDD of the model. In order to define the prior probability of the model, $p(M_j)$, one can assume that all models are given the same weight, and hence the prior can then be expressed as $p(M_j) = \frac{1}{m}$. If some models have the same explanatory power, the priors can be modified to dissuade over-parametrized models. To do so a method presented by Osiewalski and Steel (1993) [30] makes the prior inversely proportional to the number of parameters.

The article by Makiela and Osiewalski (2018) present the key results obtained through this method.

2.4 Stochastic Non-Parametric Envelopment of Data

In the previous sections, the pros and cons of DEA and SFA have been presented. Several researchers have attempted to combine the advantages of both, however the first method that truly merge the two methods in a generalized way was presented by Kuosmanen (2006) [21]. In his paper the author present a two stage model that associate the piece-wise linear frontier of DEA and the two-part stochastic residual of SFA.

To rapidly come back of the main limitations of both original models, one can indicate that the main problem of the DEA is the lack of robustness with regards to outliers. If one firm operates significantly more different than its peers, the frontier will be shifted to include the firm in the technology set. On the other hand, the main problem with the SFA method is the lack of flexibility of the functional form. In the parametric models, the functional form needs to be assumed beforehand, making the results rely heavily on the prior assumptions.

The method presented by Kuosmanen starts similarly as the SFA, by the definition of a production function.

$$y_k = f(x_k) + \hat{\epsilon_k} \qquad \forall k = 1, ..., K \tag{2.23}$$

where $\epsilon_k = u_k + v_k$ is the estimation of the residual composed of noise and inefficiency. The difference with SFA comes from the fact that the functional form $f(x_k)$ is not defined *a priori*. It is rather computed by an optimization model and a set of inequalities constraints. Before dwelving into the definition of these inequalities, let's define the characteristics of the production frontier. It is assumed that such a function must be monotonically increasing and be concave. The set of all function satisfying these assumptions is called F^2.

The StoNED model uses a non-parametric regression technique called *concave non-*

parametric least squares. The method is close to the one of the ordinary least squares, as the objective function is to minimize the squared residuals, however it differs in the sense that each observation can have its own slope and intercepts.

More formally, the optimization model can be written as

$$\begin{array}{lllll} \min\limits_{\alpha,\beta,\hat{\epsilon}} & \sum_{k=1}^{n} \hat{\epsilon}_k^2 & & & \\ s.t. & & & & \\ y_k & = & \alpha_k + \beta_k x_k + \hat{\epsilon_k} & \forall k = 1, ..., K \\ y_h & \leq & \alpha_k + \beta_k x_h + \hat{\epsilon_h} & \forall k, h = 1, ..., K \\ \beta_k & \geq & 0 & \forall k = 1, ..., K \end{array} \tag{2.24}$$

This model compute K slope vectors β_k, one for each observation. The second constraint uses the Afriat inequalities to impose concavity. The third constraint imposes monotonicity as the slope of each inequality, β_k must be positive. By taking the most inner part of each linear function, an approximation of the functional form can be determined. This model comes close to the DEA method. Indeed, the main difference comes from the fact that the error is not one sided. The term $\hat{\epsilon_k}$ can be either positive or negative in the StoNED model. It is interesting to see that if the constraint $\hat{\epsilon_k} \leq 0$ is imposed, the inequalities constraints leads to a functional form similar to the one which would be obtained using a DEA model following the VRS assumption.

After the functional form of the production function has been specified, and we have estimations for the error terms $\hat{\epsilon_k}$, we can use the same method as in the SFA approach to: (1) Find estimations for the variances of the components of the error terms $\hat{\sigma_u^2}$ and $\hat{\sigma_v^2}$, (2) Use the conditional estimator to find the inefficiency term.

While in the SFA section we used the method of the log-likelihood to estimate the variances of the error terms $v_k and u_k$, Kuoskamen uses the method of moment that allows for more computational ease.

2.4.1 Method of moments

The second and third moments can be approximated using the following estimators:

$$\begin{aligned} m_2 &= \sum_{k=1}^{K} (\hat{\epsilon_k} - \hat{E}(\epsilon_k))^2 / K \\ m_3 &= \sum_{k=1}^{K} (\hat{\epsilon_k} - \hat{E}(\epsilon_k))^3 / K \end{aligned} \tag{2.25}$$

By comparing these estimators to the true moments, it is possible to obtain estimators for the variances σ_u^2 and σ_v^2.

$$\begin{aligned} \mu_2 &= \left[\frac{\pi - 2}{\pi}\right] \sigma_u^2 + \sigma_v^2 \\ \mu_3 &= \left(\sqrt{\frac{2}{\pi}}\right) \left[1 - \frac{4}{\pi}\right] \sigma_u^3 \end{aligned} \tag{2.26}$$

Once the variances have been estimated, the same method as in the SFA approach is used to estimate the inefficiency term u. Two auxiliary variables are defined, $\mu_* = -\hat{\epsilon_k}\sigma_u^2/(\sigma_u^2 + \sigma_v^2)$, the mean of the conditional distribution of u_k, and $\sigma_*^2 = \sigma_u^2\sigma_v^2/(\sigma_u^2 + \sigma_v^2)$, representing the variance of the same distribution. The conditional estimation for u_k is the same as for the SFA, defined as

$$E(u_k|\hat{\epsilon}_k) = \mu_* + \sigma_* \left[\frac{\phi(-\mu_*/\sigma_*)}{1 - \Phi(-\mu_*/\sigma_*)}\right] \tag{2.27}$$

Another method to estimate the variances of u and v is the log-likelihood estimator. While being more computationally intensive, it can still be useful to apply it and compare the difference in outcomes in regards to the method of moments. The same parameters λ and σ as defined by Aigner et al. (1977) [7] can be used to express the log-likelihood function of the stoned model, the function takes the following form:

$$\ln L(y|\sigma^2, \lambda) = \frac{n}{2} \ln(2/\pi) - n \ln \sigma + \sum_{i=1}^{n} \ln \Phi \left[\frac{-\hat{\epsilon}_i \lambda}{\sigma}\right] - \frac{1}{2\sigma^2} \sum_{i=1}^{n} \hat{\epsilon}_i^2 \tag{2.28}$$

where $\sigma^2 = \sigma_u^2 + \sigma_v^2$ and $\lambda = \sigma_u/\sigma_v$.

2.4.2 Application of StoNED model in Electricity Distribution

Kuosmanen applies its semi-nonparametric model to the Finnish regulatory system[22]. It is worth noting that the finnish government started using this model in 2012.

The study used one input and three output variables. The input was the total expenditure (TOTEX). It is comprised of three parts: (1) capital expenditures (CAPEX), (2) operational expenditures (OPEX) and (3) the estimated cost of interruptions. This last part is estimated according to the total number and the total duration of the interruptions. The three output variables where (1) weighted amount of energy sent through the network, this quantity is weighted according to the average cost of transmission, as the voltage decrease, it is marginaly more expensive to transmit energy, so lower voltages have a higher weight in the variable, (2) the total length of the network in kilometer and (3) the total number of customer plugged to the network. The values of those variables of the 4 years before the regulation process.

An other variable that is not an input nor an output is added, z is a contextual variable that account for the heterogeneity between the firms. It account principally for the proportion of underground cable. This types of cables are more expensive to implement and manage.

The cost frontier model used in the study defined as follow:

$$x_k = C(y_k) \times \exp\left(\delta z_k + u_k + v_k\right) \tag{2.29}$$

where $C(y_k)$ is the cost function according to the output y_k. The parameter delta account for the weight of the heterogeneity among firms and in particular for the weight of the proportion of underground cables, z_k.

The assumptions regarding the cost function are the same as for the StoNED model presented earlier, that is (1) C is monotonic, (2) C is convex and (3) return to scales of C

are constant.

The cost function is linearized by applying a log function on both sides, and the first step of the StoNED approach is applied, the square of error term is minimized and the constrained nonlinear least squares (CNLS) constraints are applied.

$$\begin{array}{rcll} \min\limits_{\gamma,\beta,\delta,\epsilon} & \sum_{k=1}^{n} \epsilon_k^2 & & \\ s.t. & & & \\ \ln(x_k) & = & \ln\gamma_k + \delta z_k + \epsilon_k & \forall k = 1, ...K \\ \gamma_k & = & \beta_k y_k & \forall k = 1, ..., K \\ \gamma_k & \geq & \beta_h y_k & \forall h, k = 1, ..., K \\ \beta_k & \geq & 0 & \forall k = 1, ...K \end{array} \tag{2.30}$$

Once the parameters γ_k, β_k, δ and ϵ_k have been estimated, the method of moments is used to estimate the variances of u_k and v_k. Then using the conditional expectancy of u_k knowing $epsilon_k$, using the mean and variance of the conditional distribution, we can have an unbiased estimator of u_k. As stated in a previous section, while unbiased this estimator is inconsistent.

As the inputs used in this application are the costs, the cost efficiency can be calculated using the same technique as in the SFA. By taking the exponent of $\hat{u}_k$, $\exp u_k$ we obtain a value comprised between 0 and 1 representing the cost efficiency. The inefficiency term here is not preceded by a negative operator, as we are calculating the output efficiency. A positive value of u_k means that less outputs have been produced for the same level of input.

2.4.3 Difference with other models

Kuoskamen lists the main differences to the DEA program:

- In contrast to DEA, the StoNED approach solves the residual $\hat{\epsilon_k}$ simultaneously,

whereas DEA solves it independently for each observation, as there is one linear program problem for each firm.

- The nature of the inefficiency term is also a major difference between the two approaches. In the DEA method, the technical efficiency is multiplicative and expressed as a percentage relative to the production frontier, whereas in the StoNED model, the inefficiency term u_k is on an absolute scale and represent the absolute variation from the functional form not accounting for the random noise.

- The last main difference raised by Kuoskamen is that while the DEA functional form of the production frontier is implicit, with the StoNED method the sub-gradients are explicitly defined by the optimization model. They correspond to the parts of the Afriat inequalities that constrain the most the estimated production frontier.

Chapter 3

Empirical Testing

The aim of this chapter will be to apply each method descried above using the R software. First, the data that will be used will be introduced. Then, the DEA and SFA approaches will be performed. These two approaches have already received sufficient attention and researchers have already come up with R packages to solve them. This thesis will use the package "Benchmarking" developed by Bogetoft and Otto. A R package for the StoNED method hasn't been developed yet, so the first task in this regard will be to develop a set of routines that will solve the StoNED model in R. The algorithms that will be used are retrieved from Johnson and Kuosmanen (2014). This chapter will mainly be focused on the application of the models, the results obtained will be discussed in the next chapter.

3.1 Model And Data Pre-Processing

The data that is going to be used is sourced from the Swedish DSO's. It is gathered on a time span of seven years, from 2000 to 2006. However, the application of the different models on time series is not part of the scope of this project, and thus will be omitted. The empirical testing will be thus conducted separately for each of the seven year we have data on.

The different metrics that were retrieved from the DSO's cover different aspects of the companies: they range from the costs to the different operational metrics specific to the electricity distribution sector. The most important variables will be presented bellow, but a full table of the variables provided by the data set can be found in the appendix.

The data set contained some missing values. The strategy concerning those has been to remove the observations containing missing values for the variables that are used in the model. This strategy is acceptable as the updated data sets all contain approximately 200 observations.

3.1.1 Outlier Management

Some units of the dataset are outliers. There exists different ways to manage the outliers. The one that was used here relied on the superefficiency values computed using the sdea function of the Benchmarking package. This technique is derived from the one used by Agrell and Niknazar (2013) [5].

Once the superefficiency have been computed, the values that were more or less than 2 times the standard deviation from the median were removed from the dataset. This techniques permitted to have cleaner results and remove observations that could contain errors.

3.1.2 Variable selection

The choice of variables is an essential matter for the regulator. A badly specified model could lead to unfair revenue allocation, which in turn would lead the regulated operators to challenge the computed studies. A set of models are defined in Agrell and Bogetoft (2005) [4] and Agrell and Niknazar (2013) [5]. The authors distinguish two types of model, depending on the time span on wants to cover. The long-term model (LR) takes into account the different types of costs, from OPEX to CAPEX, while the short-term model (SR) only takes into account the costs that can be managed on the span of one

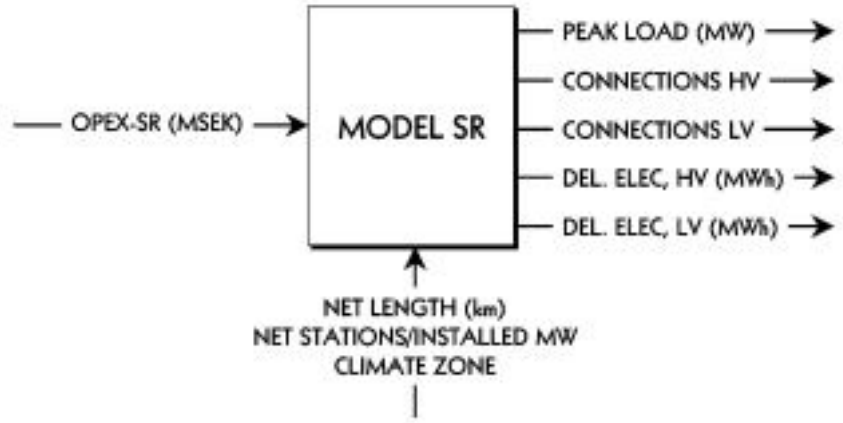

Figure 3.1: Short Term Model

Source: Agrell and Bogetoft, 2005

year. It thus exclude the CAPEX, as these expenditures are decided on longer term. The net losses are also excluded.

While the inputs are quite different between the two models, the outputs remain unchanged. the main outputs are the number of connections, the energy delivered and the peak load.

For this master thesis, we will focus on the long-term model, where the inputs will not be the three variables presented in figure 3.2, but rather the total expenditure, as this metric englobes the three inputs of the presented model. The outputs on the other hand will remain unchanged.

Multiple outputs

The models presented above have a particularity that was not discussed in the previous chapter. Indeed, these models have one or more input, and five outputs. The DEA model is designed to work with multiple input and outputs, however the SFA and StoNED approaches won't be able to solve these models with the formulations we have presented above. The use of distance functions will thus be required. The variables will have to be reworked in order to have one output and several inputs. The method that will be

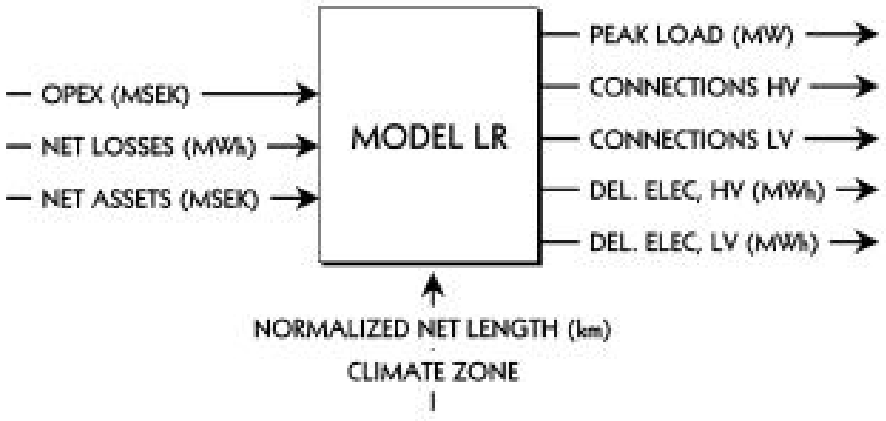

Figure 3.2: Long Term Model

Source: Agrell and Bogetoft, 2005

used is presented in Bogetoft and Otto (2011) [10]. The specificity of the method will be presented in the SFA section.

3.2 DEA

The first model to be applied is the DEA. This model has as a core advantage that it handles the use of multiple inputs and outputs. The function "dea" of the package "Benchmarking" is thus fully capable of handling the model that will be applied.

The model has been run on the seven years, the different results obtained will be analysed and compared. The first application will be using the varying return to scale assumption, but an analysis will be conducted on the different assumptions later. The model is input-oriented, as it makes more sense for a DSO to provide the same level of service at a lesser cost. Input-oriented DEA evaluates by which factor the input could be cut while providing the same level of output.

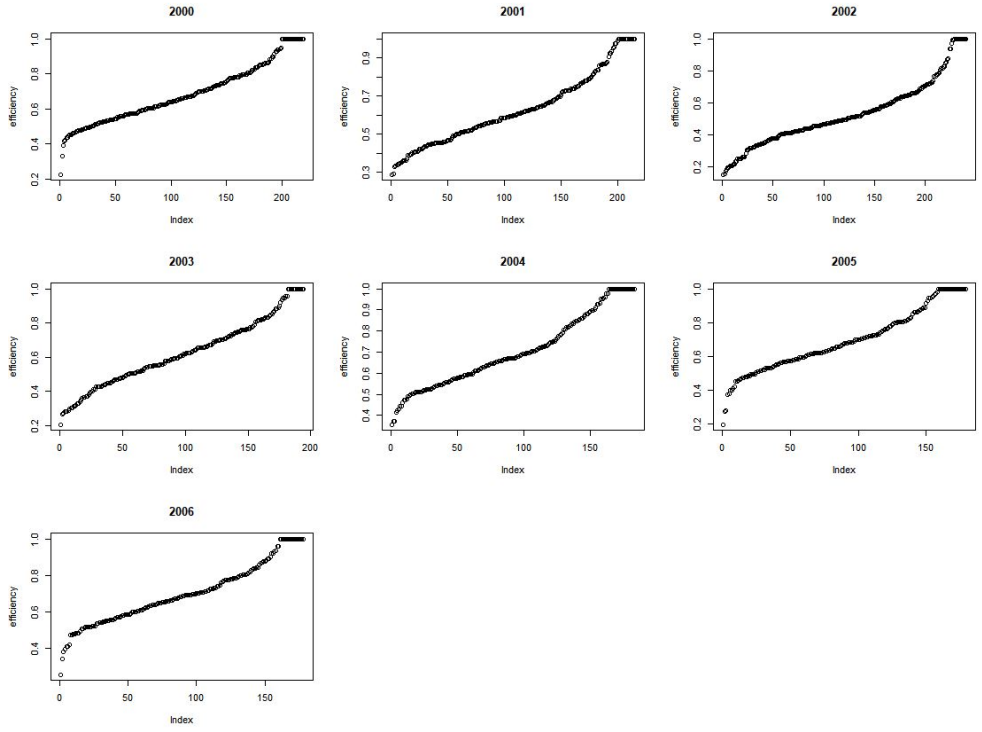

Figure 3.3: Sorted Efficiencies Per Year, varying return to scale DEA

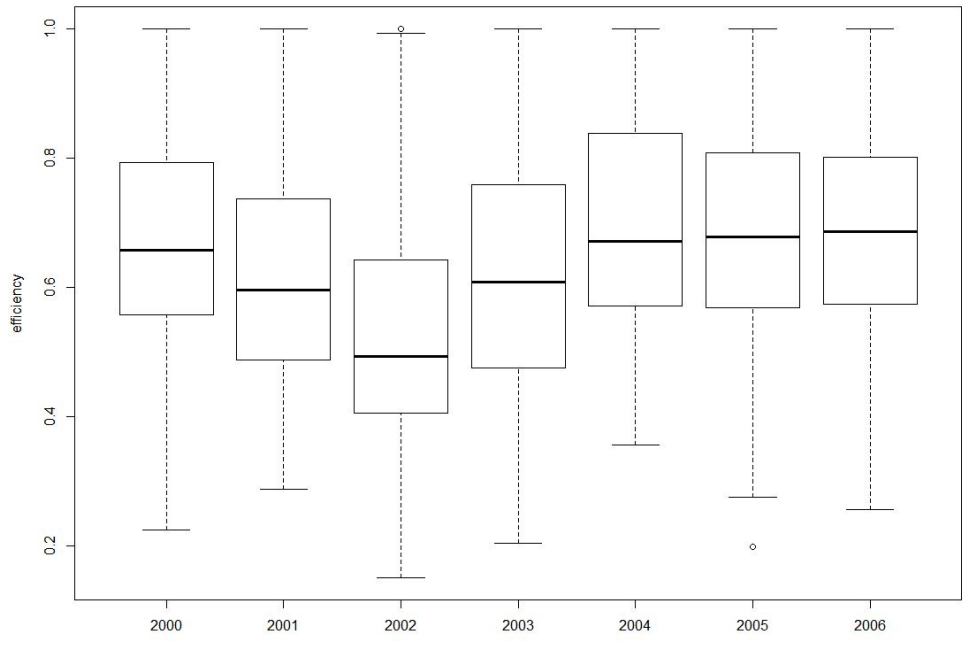

Figure 3.4: Boxplots of Efficiencies, varying return to scale DEA

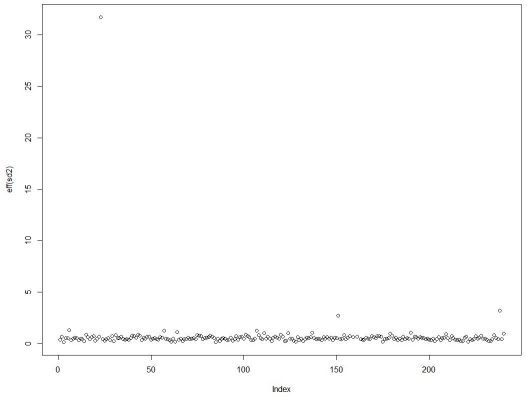

Figure 3.5: Super efficiency of 2002's DSO's, VRS assumption

3.2.1 Results

The graphs of 3.3 and 3.4 represent the efficiencies we obtained for each year. We can see that the efficiencies range mostly form 0.2/0.3 to 1. One can observe in the boxplots that the year 2002 presents a smaller median efficiency than the other years. This can be explained by the presence of an outlier during this year. In fact, the DSO number 24 presents a very low TOTEX (23 000 kr), for an amount of delivered low-voltage energy of 14 361 MWh. These quantities make the DSO very efficient compared to its competitors. The frontier is thus pushed by this particularly well performing DSO, hence the distance of many other DSO's from the frontier is increased, resulting in a lower computed efficiency.

The distance of this firm from its direct competitors can be calculated using the super efficiency. With the super efficiency, the firm we want to estimate the efficiency is omitted in the estimation of the technology set. The super input efficiency can thus be greater than one if the observed DSO is outside of the set. The magnitude of the super efficiency can help grasp the relative efficiency of firms that both have an efficiency of 1 with a standard DEA. The function "sdea" of the "Benchmarking" package is used to compute the super efficiency.

What can directly be observed in 3.5 is that the DSO number 24 has a super efficiency dwarfing all the others. This would imply that it is much more efficient than all the other observations that also have an efficiency of 1 under normal conditions.

Remember that the particularity of DEA is that it interprets all deviations from the frontier as inefficiency, even though the presence of outliers in the observations could be due to noise in the data.

3.2.2 Return to scale assumptions

The "dea" function allows us to modify the return to scale assumptions. Further models have thus been run in order to observe the difference between return to scale assumptions. The resulting graphs of these applications can be found in the appendix.

What can be first observed, is that the assumption on return to scale can really impact the computed efficiency. When the CRS assumption is considered, the technology set is larger than with the other assumptions, resulting in equal or lower efficiency for the analysed DSO's. For most of the years observed, the result is a small downward shift of the efficiency of most firms. However the impact on the year of 2002 is much larger. Because of the presence of the DSO number 24, almost all of the firms have seen their efficiencies shrink. Because of the small size of this observation (TOTEX is 23 000 kr, while the mean of 2002 is 64 million kr), the impact on other firms was contained under the VRS assumption. But under CRS, the frontier is not only shifted for the DSO that have similar scales of inputs and outputs, but for the whole set of DSO's. The effect of this outlier is thus expanded to the whole sector. The effect can be clearly observed on A.3. The median decreased up to around 0.25 and the third quartile is only slightly above 0.4.

The similar impact can be seen with the IRS assumption. The increasing (or non-decreasing) return to scale assumption make the RTS constant when a DSO is scaled up. Because the DSO number 24 is very small, the impact is thus the same than with the CRS assumption.

The free disposability hull return to scale has the smallest technology set, and thus is the most tolerant towards the observed firms. With this assumption, the median efficiency is almost always above 0.8.

3.3 SFA

The second method that will be applied is the SFA. As explained in the previous chapter, the SFA has the particularity that it integrates the concept of noise in the data. The deviation form the frontier is now not only due to inefficiency, but also due to a random term.

The *Benchmarking* package is also comprised of a function that permits to run SFA models. However, the R method *sfa* only allows for one output. In our case, we have one input and several outputs.

3.3.1 Estimable distance function

In order to allow the existing R method to run the model we are applying, we will make use of *estimable stochastic distance function* which is presented in Bogetoft and Otto (2011)[10].

Consider the input and output distances defined by equation 2.3. One of the properties of the Shepard distance function is homogeneous of degree 1 in x for the input, and in y for the output, which mathematically translates as:

$$\begin{aligned} D_i(tx, y) &= tD_i(x, y) \\ D_o(x, ty) &= tD_o(x, y) \end{aligned} \tag{3.1}$$

We can defined the inefficiency term u as being the distance an observation is form the frontier, hence for the input distance case:

$$D_i(x, y) = e^u \tag{3.2}$$

Due to the homogeneity of the function we have

$$x_m D_i(\frac{x}{x_m}, y) = D_i(x, y) \tag{3.3}$$

where $\frac{x}{x_m}$ is interpreted as $\left(\frac{x_1}{x_m}, ..., \frac{x_{m-1}}{x_m}\right)$. Then if we take the logarithms of this equation we get

$$log(x_m) + log(D_i(\frac{x}{x_m}, y)) = log(D_i(x, y)) \tag{3.4}$$

By equation 3.2, we can modify the previous equation as

$$-log(x_m) = log(D_i(\frac{x}{x_m}, y)) - u \tag{3.5}$$

The last equation has almost the same functional form than the stochastic production frontier model form equation 2.11. By adding a stochastic element v to the right hand side of the equation we can obtain the exact formulation. The new formulation is

$$-log(x_m) = log(D_i(\frac{x}{x_m}, y)) + v - u \tag{3.6}$$

This formulation can now be used with the *sfa* method of the *Benchmarking* package to obtain the technical efficiencies.

Note that even though the previous development was made using only the input distance function, the same reasoning can be achieved for the output distance function. The main difference being that now the new formulation is expressed as

$$log(y_n) = -log(D_o(x, \frac{y_n}{y})) + v - u \tag{3.7}$$

In order to apply the method presented above, the inputs and outputs needed to be modified. The new X and Y parameters of the *sfa* method are:

- Input Distance Function:

X	yEnergy.del.lv	yEnergy.del.hv	yPeakload	yConnections.lv	yConnections.hv
Y	dTotex				

- Output distance function:

X	$\frac{yEnergy.del.hv}{yEnergy.del.lv}$	$\frac{yPeakload}{yEnergy.del.lv}$	$\frac{yConnections.lv}{yEnergy.del.lv}$	$\frac{yConnections.hv}{yEnergy.del.lv}$	dTotex
Y	yEnergy.del.lv				

3.3.2 Results

The next step of the analysis was to run the two SFA models. Indeed, for this step both the input and output distance function will be applied.

Input distance function

The input distance function was the first to be run, and the results obtained can be found in Fig. 3.6.

Table 3.1 groups the value of lambda for each year. The lambda for the 2001 model is negative, which is the reason why the efficiencies for this year are not correct. Furthermore, we can see that the lambda of 2002 and 2003 are close to zero. It would imply that most of the deviation from the production function is due to the random noise and not the efficiency. The proportion of deviation explained by the inefficiency can be calculated by equation 2.14. These values have been calculated and have also been reported on table 3.1. The value for 2001 has not been computed because it is not relevant to calculate the proportion if the lambda is negative. The values for 2002 and 2003 on the other hand show that all the deviation form the production function is due to the random noise. This explains the fact the efficiencies on the histograms 3.6 of 2002 and 2003 are concentrated close to 1.

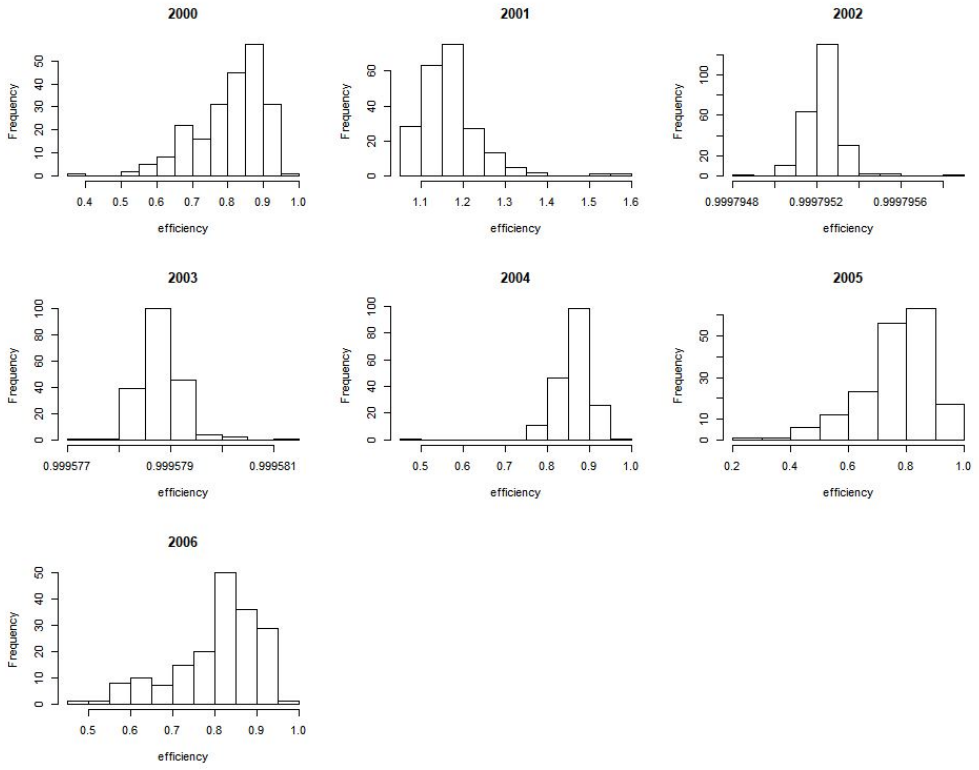

Figure 3.6: Histograms of SFA efficiencies

Further analysis can be done on the parameters of the log-likelihood function to see if the estimated lambda, variances and beta parameters are statistically significant. What resulted from this analysis is that the values of lambda for the year 2002 and 2003 both have a p-value of 0.99, which makes them unreliable. We can thus conclude that the model could not find any inefficiency values for the years 2002 and 2003.

Year	2000	2001	2002	2003	2004	2005	2006
Lambda	1.871	-0.885	0.001	0.002	0.863	2.378	1.988
Prop. Dev.	77%	/	0%	0%	43%	85%	80%

Table 3.1: Values of lambda

Output distance function

The next step was to compute the efficiencies for the SFA output estimable distance function. The output technical efficiency was here considered, because of the poor significance

of some parameters of the input estimable distance function.

The efficiencies computed can be found in figure 3.7. The results for 2002 and 2003 are better as now the proportion of deviation due to inefficiency is greater than 0, these values can be found in table 3.2. It is clear that the model for the output distance function is more reliable than for the inputs distance function.

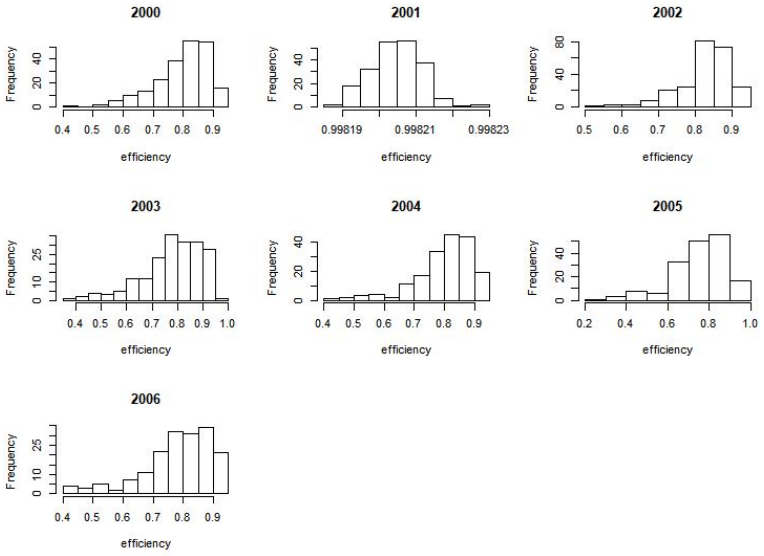

Figure 3.7: Histogram of efficiency for SFA output function

Table 3.2: Values of lambda

Year	2000	2001	2002	2003	2004	2005	2006
Lambda	2.087	0.008	1.425	3.023	2.431	2.769	2.186
Prop. Dev.	81%	0%	67%	90%	85%	88%	83%

3.4 StoNED

One of the main goals of this master thesis was to implement a method to solve the StoNED model in R. Papers such as Johnson and Kuosmanen (2014) [19] have already shown methods to solve it in matlab or GAMS, but R being an open-source software, a new package implementing the StoNED method could make the method more accessible.

The StoNED approach comprises two step: (1) the estimation of the CNLS piece-wise frontier and (2) the interpretation of residuals using the method of moments or the log-likelihood estimator.

3.4.1 Implementing the model

First, the CNLS needed to be solved. R has numerous packages to solve optimization problems. The CVXR package was used due to its flexibility and ease of interpretation. This package is an extension for R of the matlab's package CVX, which is used by Johnson and Kuosmanen in the article mentioned above.

The quadratic optimization problem implemented is the following:

$$\begin{array}{lllll} \min\limits_{\alpha,\beta,\hat{\epsilon}} & \sum_{k=1}^{n} \hat{\epsilon}_k^2 & & & \\ s.t. & & & & \\ y_k & = & \alpha_k + \beta_k x_k + \hat{\epsilon_k} & \forall k = 1, ..., K \\ y_h & \leq & \alpha_k + \beta_k x_h + \hat{\epsilon_h} & \forall k, h = 1, ..., K \\ \beta_k & \geq & 0 & \forall k = 1, ..., K \end{array} \tag{3.8}$$

According to Lee et al. (2011) [28]: "In general the parameters α and β estimated using CNLS are non-unique; however the fitted values, $\hat{y}_i = \hat{\alpha}_i + \hat{\beta}_i x_i$, are unique." Hence, in order to correctly estimate the lower concave envelope of the CNLS regression, the following model is used.

$$\begin{array}{rl} \min\limits_{\alpha,\beta} & \alpha_i + \beta_i * x_i \\ s.t. & \\ \alpha_i + \beta_i * x_i >= \hat{y}_i & \end{array} \tag{3.9}$$

This models requires as parameters the fitted values $\hat{y}_i$ of the previous model. Note that here the residuals are not affected, as they are the difference between $\hat{y}_i$ and y_i, which are

two parameters in this model. The goal is thus to reestimate the lower concave envelope.

Once the residuals of the CNLS model have been correctly estimated, they can be applied as in the SFA method to derive the standard deviations of the inefficiency and random noise terms. The SFA model relied on the use of a log-likelihood estimator. While this method still can be used here, Kuosmanen introduces the method of moments to find the variance of random noise σ_v^2 and the variance of inefficiency σ_u^2. This method relies on the two set of equations 2.25 and 2.26 to find the estimated values.

The R routine for the StoNED model has been implemented with both method, so that they can be compared in different situations. The log-likelihood estimator is the one devised by Kuosmanen (2006), and takes the form of equation 2.28.

Once the variances of the distribution of u and v have been estimated, the method of the conditional mean of Jondrow et al. (1982) [20] has been applied to find estimations of the inefficiency term u.

It can also happen that the conditional expectation gives us inefficiency values below zero or above 1. This happens if the errors terms found by the CNLS method are too extreme. The solution that has been applied here is to bound the efficiency values between 0 and 1.

3.4.2 Results

The methods presented above have thus been applied to the data, figure 3.8 shows the distribution of the computed efficiencies per year. Note that similar to the efficiency found using the SFA method, most of the efficiencies have a value close to 0.8/09, with a small tail which can extend up to 0 in some cases. Some efficiencies are equal to zero, but this is due to the bounding explained in the previous section.

The bounding of the inefficiency term can happen for various reason. One of them, which is applied for the year 2005, is that there is a outlier. Indeed, when analysing the histogram of the error values computed by the CNLS model, figure 3.9, we can observe that

the error are negatively skewed. This skewness is caused by a firm having particularly bad records, which results in the efficiencies of the other firms being positively shifted towards an efficiency of 1. Figure A.17 shows us that particular shift as we observe that a lot of efficiencies are equal to 1.

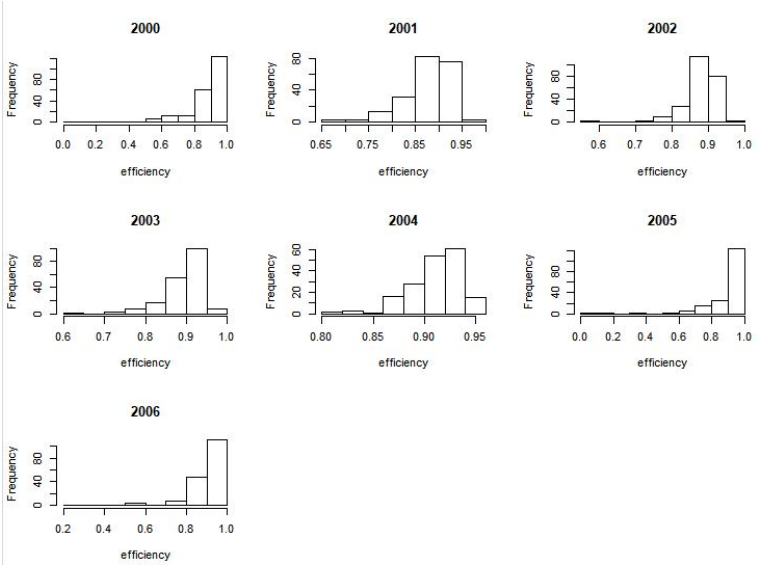

Figure 3.8: Histogram of efficiency for StoNED model

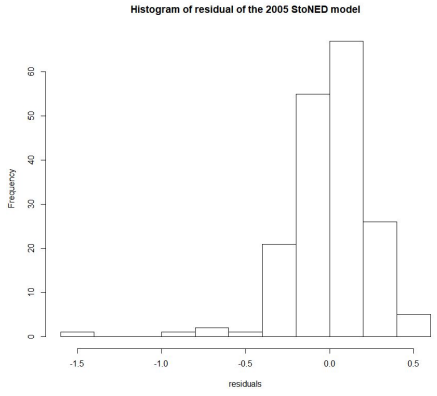

Figure 3.9: Histogram of CNLS residuals, 2005

The validity of the CNLS parameters an be estimated using the R^2. The formula used to assess the significance of the models is

$$R^2 = 1 - \frac{\sum_{i=1}^{K} e_i^2}{\sum_{i=1}^{K} (y_i - \overline{y})} \tag{3.10}$$

3.4.3 Log transformation

As stated above, in some situation the Afriat inequalities are particularly difficult to solve. One of the reason is developed in the previous section. However sorting the data or making it follow a particular structure is sometimes not sufficient.

In fact, let's take the first two constraints of model 3.8:

$$\begin{aligned} y_k &= \alpha_k + \beta_k x_k + \hat{\epsilon_k} \quad \forall k = 1, ..., K \\ y_h &\leq \alpha_k + \beta_k x_h + \hat{\epsilon_h} \quad \forall k, h = 1, ..., K \end{aligned} \tag{3.11}$$

The first constraints defined the slope parameters, and the second imposes concavity. Unfortunately, the CNLS model might be infeasible under certain circumstances. When the variance is too great, the cnls model cannot respect constraint 1 and 2 simultenaously, because by increasing the error term to honor constraint 2, it has to decrease the parameters values to respect constraint 1, which cancels the effect of the first increase of the error term and makes the model infeasible.

The fact that the CVXR solver cannot find a solution can seem quite odd as a potential solution is to make all the parameters equal to 0, and the error term equal to the corresponding output. While this solution is feasible, it has no interest for us as it means that the efficient frontier is to produce nothing for all input levels.

A partial solution that works for all the data sets that were tested is to take the logarithm values of the inputs and outputs. Taking the logarithms reduce the variance of the observations and help the solver to find an appropriate solution.

Chapter 4

Analysis

In this section, the results we obtained in the previous section will be analysed against each other. The similarities as well as the differences between the three models will be discussed.

4.1 Descriptive Analysis Table

The inefficiency values obtained using the different models will be discussed here. For simplicity measure, only the data of the most recent year, 2006, will be analysed. A more complete picture of the results obtained can be found in the appendix.

What can be observed is that there is a separation between the DEA and SFA model and the StoNED one. Concerning the last one, the mean is much higher and the standard

Model	Mean	Std	Min	Max	1st Quart.	2nd Quart.	3rd Quart.
DEA	0.703	0.160	0.257	1	0.583	0.691	0.803
SFA	0.778	0.116	0.415	0.942	0.732	0.801	0.863
StoNED	0.897	0.095	0.273	0.979	0.871	0.926	0.951

Table 4.1: Decriptive Analysis

deviation much smaller. The inefficiency values are more concentrated. However, the minimum value for the StoNED model is low, which implies that there is a large tail in the distribution of the inefficiencies. What could be thus inferred is that the StoNED model punishes a small percentile of units that performs very badly.

4.2 Correlations

The efficiencies obtained through the three models have been confronted against each other in order to see if there was indeed a correlation between the different models. Figure A.18 the model correlation of each years.

What can be observed is that the DEA efficiencies are more spread out than their SFA and StoNED counterparts. This is in fact confirmed by looking at table 4.2 which contains the variances of each model. The variances of the DEA model are always larger compared to the two other models. This can also be oberved by looking at the boxplots of the different models, figure 4.1 shows three different boxplots representing the three different models for the year 2000. What can be seen is that the efficiencies of the DEA method are more spread out, while the efficiencies computed using the StoNED method are more concentrated around 0.9. The StoNED model also has the particularity of having a tail. 75% of the observations are comprised between 0.7 and 1, but the rest of the firms have an efficiency that can extend down to 0.

The volatility of the DEA's efficiencies can also be explained by the flexibility of DEA approach. Because this model attributes all the deviation form the frontier to the inefficiency term, the presence of outliers has a large impact on the computed efficiencies. The random noise term thus reduce the impact of these outliers. As a result, we can see that the efficiencies found through the SFA and StoNED models are more consistent and less likely to vary when the data has extreme outliers.

Table 4.3 shows the correlation values obtained by comparing the different models against themselves. The correlations between DEA and SFA are lower than the other cor-

Model	2000	2001	2002	2003	2004	2005	2006
DEA	0.027	0.034	0.040	0.038	0.029	0.032	0.027
SFA	0.017	0.012	0.008	0.019	0.016	0.020	0.014
StoNED	0.016	0.003	0.004	0.011	0.012	0.022	0.009

Table 4.2: Variances of inefficiencies

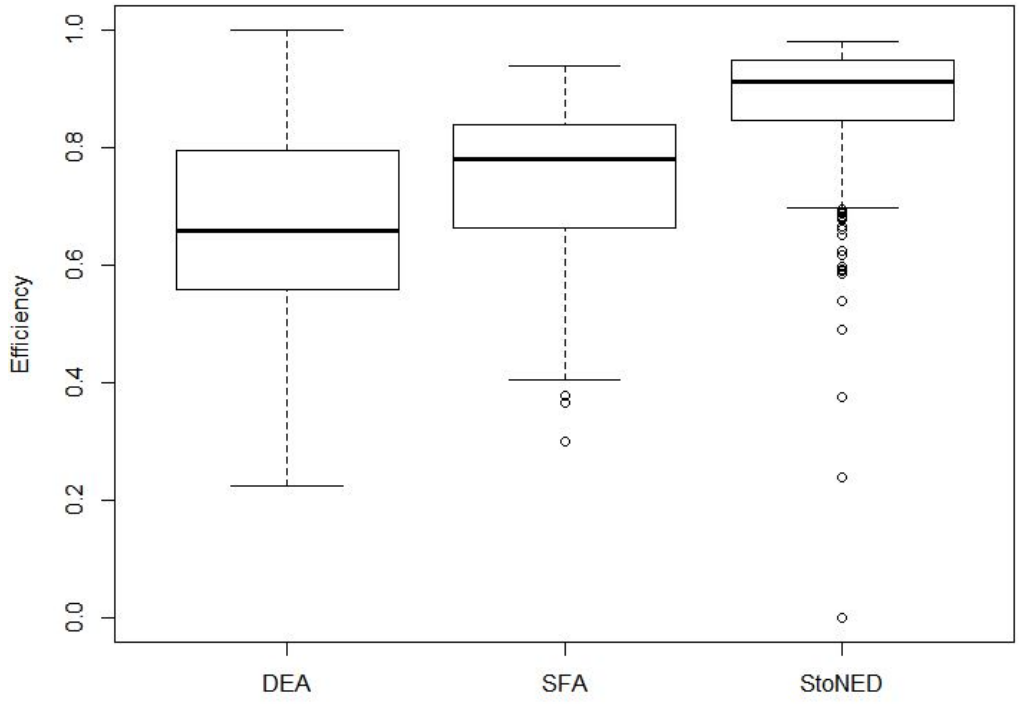

Figure 4.1: Efficiencies of 2000

Models	2000	2001	2002	2003	2004	2005	2006
DEA/SFA	0.503	0.442	0.467	0.343	0.410	0.447	0.407
SFA/StoNED	0.673	0.690	0.587	0.544	0.624	0.648	0.641
DEA/StoNED	0.661	0.627	0.837	0.654	0.674	0.653	0.624

Table 4.3: Correlations between models

relations, they range between 0.35 and 0.5. On the other hand, the efficiencies computed with the StoNED method seems to have a high correlation with both the DEA and SFA efficiencies.

The R-squared values estimated by equation 3.10 are usually above 0.9. However these values are computed using the parameters estimated by the CNLS model 3.9. The lower concave envelope found is thus optimal.

The outliers also have an impact on the validity of the model. As an example, the StoNED model has been run for the year 2003, but only on one output. When all the firms are included in the model, the method of moment gives us a positive third moment, which indicates that the CNLS residuals are skewed in the wrong direction. The third moment is equal to 0.152. If we exclude the most extreme value, the firm 23, the third moment is much closer to 0, equal to 0.0049.

Overall the structure of the efficiencies found through the StoNED model seem to correspond with the one of the EA and SFA models. The correlations involvnig the StoNED model are higher than the correlation between the two more used models.

4.3 Model Limitations

4.3.1 Data Validity

It can happen that the CNLS model is unable to solve the Afriat inequalities. One of the first or second constraint of the model 3.8 is violated, and both cannot be solved together. This problem arises particularly when the data is randomly generated. This

implies that the data that will be tested must obey a particular structure. When random data is generated and used as inputs and outputs for the model, the CNLS model gives very bad results and the R^2 is very low (usually around 0.1 or 0.2). The hypothesis is thus to try and see if the model behave differently when we add a structure to the data.

The first structure that will be tested is thus to sort the both the inputs and outputs in increasing order. By sorting the data in this fashion, what can be observed is that the CNLS model can find a much better solution compared to the unsorted one. The difference has been tested multiple times using the code illustrated in figure A.19 and the results where always significant.

But only sorting the data in increasing order is a pretty strong hypothesis. To generate data in a more natural way, one can use the method used in Kuosmanen (2006).

The method goes as follow: The input values are randomly generated, in our case, they follow the uniform distribution. A true efficiency frontier is arbitrarily chosen, for example $eff\_frontier = \ln(input) + 2$. The noise and inefficiency terms are separately generated, with a mean equal to 0 and an arbitrarily determined standard deviation. Let's note that the inefficiency term needs to be bounded at 0 to be positive. The simulated output values are then computed by subtracting the random terms from the true efficiency function.

In our case, 200 observations have been generated, and the error terms follow the following distribution, $u_i = |N_{iid}(0, 0.6^2)|$ and $v_i = N(0, 0.3^2)$. Graph 4.2 shows the simulated output in correspondence with the simulated inputs. The red line represents the true efficiency frontier.

The cnls model has been run multiple times with new randomly generated data. Ideally, the estimated value of σ_u and σ_v should be respectively equal to 0.6 and 0.3. after 50 new sets of data, the mean values found by the CNLS model were 0.587 and 0.288. It can thus be concluded that the model is quite competent to find a good estimate of the standard deviation of the error terms.

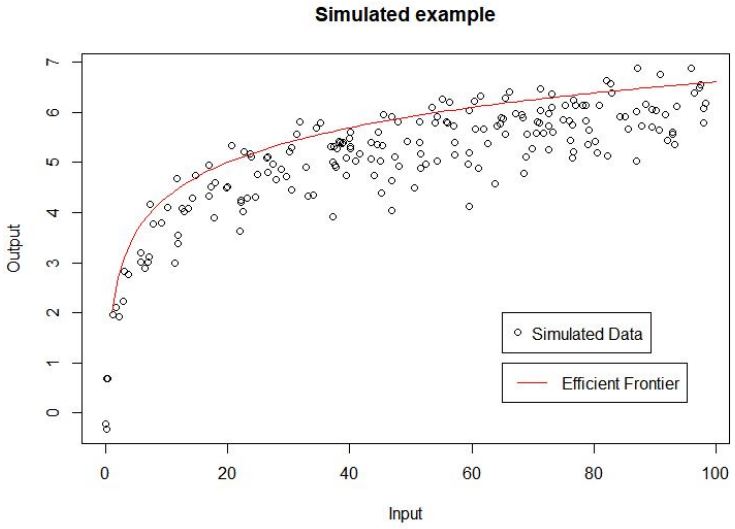

Figure 4.2: Graph of simulated inputs and outputs

4.4 Important Results

This section will summarize the key results found by the application of the different models in R. The results of the StoNED model application using randomly generated data as in the previous section will be assessed.

First, the validity of the implemented StoNED model has been assessed using randomly generated data. The model has been run multiple times using generated inputs and outputs. These inputs follow the uniform distribution, and are bounded between 0 and 10000. The efficiency frontier is generated as described in the previous section, specifically $eff_frontier = log(1 : (maxobs)) + 2$. A random standard deviation for the noise and the inefficiency values have been set. In this case, they are respectively equal to 0.3 and 0.6. The generated outputs follow the rule : $sim_output = log(sim_input) + 2 - noise - ineff$, where $noise$ and $ineff$ are the randomly generated noise and inefficiency values for each of the observations.

The StoNED model has been tested 50 times using the following data structure, and the computed values for the standard deviation of the noise and inefficiency values have been averaged. The mean values are 0.587 and 0.288. It can thus be assumed that the

model is capable of finding good estimations of the standard deviation of the noise and inefficiency terms.

The model has then been run on the provided data of the Swedish DSO's. A summary of the results can be found in the descriptive analysis section and in the appendix A.3.

By looking at table 4.3, we can conclude that the stoned model is correlated with the SFA and DEA models. This can be seen in the graphs of appendix A.4.

Concerning the time of computation, the DEA and SFA models run almost instantly on the dataset that has been provided. However, the implementation of the StoNED model uses the CVXR package and optimisation function that take more time to compute. The time of computation varies between 5 and 10 seconds. The DEA and SFA models are thus much more efficient. A more optimised model for the StoNED model could be developed using another package for convex non-linear optimisation problems.

Conclusion

The objective of this master thesis was to implement the StoNED model in R and see if the results could be interpreted and relied upon by the energy network regulators.

The results obtained with the StoNED model seemed to correlate with the two other models. While often these results had a more concentrated distribution around the higher values with a tail containing the least efficient units, the correlation with the efficiency factors of the two other models are significant. This implies that this method could be used for the benchmarking of DSO's as the results seemed to respect the real performance of the unit.

The main advantage of the StoNED model against the two other is that it combines the two strenghts of these models: a stochastic factor to account for the noise in the data and the fact that the functional form of the efficiency frontier is not predefined.

However, for a new Benchmarking model to be accepted by the European regulators, it must be extensively tested, as the efficiency scores computed using a particular model will have important real-life consequences for the companies operating the network and the consumer's energy bill.

Additionally, the grid structure is rapidly changing, the transition to a greener energy generation will furthermore change the network structure. The benchmarking models will need to adapt to the changes of the grid structure. Thus, the analyses made in this master thesis also need to be retested using more recent data.

Bibliography

[1] Handbook of Applied Econometrics Volume II: Microeconomics. In M. Hashem Pesaran and Peter Schmidt, editors, *Handbook of Applied Econometrics Volume II: Microeconomics*. Blackwell Publishing Ltd, Oxford, UK, 1999.

[2] Incentives Schemes for Regulating Distribution System Operators, including for innovation. Technical Report C17-DS-37-05, Council of European Energy Regulator, Brussels, 2018.

[3] Per J. Agrell. Incentive Regulation of Networks. Concepts, definitions and models. *Reflets et perspectives de la vie économique*, LIV(1-2):103–132, 2015.

[4] Per J. Agrell, Peter Bogetoft, and Jørgen Tind. DEA and Dynamic Yardstick Competition in Scandinavian Electricity Distribution. *Journal of Productivity Analysis*, 23(2):173–201, May 2005.

[5] Per J. Agrell and Pooria Niknazar. Structural and behavioral robustness in applied best-practice regulation. *Socio-Economic Planning Sciences*, 48(1):89–103, March 2014.

[6] Per J. Agrell and Jonas Teusch. Making the Belgian distribution system fit for the energy transition: The case for yardstick competition. *Reflets et perspectives de la vie économique*, LIV(1):157, 2015.

[7] Dennis Aigner, C.A.Knox Lovell, and Peter Schmidt. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, 6(1):21–37, July 1977.

[8] Rajiv Banker, William Cooper, Lawrence Seiford, and Joe Zhu. *Returns to scale in DEA*, pages 41–70. 08 2011.

[9] Peter Bogetoft. Dea-based yardstick competition: The optimality of best practice regulation. *Annals of Operations Research*, 73(0):277–298, Oct 1997.

[10] Peter Bogetoft and Lars Otto. *Benchmarking with DEA, SFA, and R*, volume 157 of *International Series in Operations Research & Management Science*. Springer New York, New York, NY, 2011.

[11] R. G. Chambers, Y. Chung, and R. Färe. Profit, Directional Distance Functions, and Nerlovian Efficiency. *Journal of Optimization Theory and Applications*, 98(2):351–364, August 1998.

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

[13] Xiaomei Cheng, Jonas Andersson, and Endre Bjorndal. On the Distributional Assumptions in the StoNED Model. *SSRN Electronic Journal*, 2015.

[14] Tim Coelli and Tim Coelli, editors. *An introduction to efficiency and productivity analysis*. Springer, New York, 2nd ed edition, 2005.

[15] M. J. Farrell. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120(3):39, 1957.

[16] William H. Greene. The Econometric Approach to Efficiency Analysis. In Harold O. Fried, C. A. Knox Lovell, and Shelton S. Schmidt, editors, *The Measurement of Productive Efficiency and Productivity Change*, pages 92–250. Oxford University Press, February 2008.

[17] Mark A. Jamison. *Regulation: Price Cap and Revenue Cap*, volume 3. CRC Press, Taylor and Francis, New York.

[18] Andrew L. Johnson and Timo Kuosmanen. One-stage estimation of the effects of operational conditions and practices on productive performance: asymptotically normal

and efficient, root-n consistent StoNEZD method. *Journal of Productivity Analysis*, 36(2):219–230, October 2011.

[19] Andrew L. Johnson and Timo Kuosmanen. An Introduction to CNLS and StoNED Methods for Efficiency Analysis: Economic Insights and Computational Aspects. In Subhash C. Ray, Subal C. Kumbhakar, and Pami Dua, editors, *Benchmarking for Performance Evaluation*, pages 117–186. Springer India, New Delhi, 2015.

[20] James Jondrow, Ca Lovell, Ivan Materov, and Peter Schmidt. On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics*, 19:233–238, 02 1982.

[21] Timo Kuosmanen. Stochastic Nonparametric Envelopment of Data: Combining Virtues of SFA and DEA in a Unified Framework. *SSRN Electronic Journal*, 2006.

[22] Timo Kuosmanen. Stochastic semi-nonparametric frontier estimation of electricity distribution networks: Application of the StoNED method in the Finnish regulatory model. *Energy Economics*, 34(6):2189–2199, November 2012.

[23] Timo Kuosmanen. Stochastic semi-nonparametric frontier estimation of electricity distribution networks: Application of the StoNED method in the Finnish regulatory model. *Energy Economics*, 34(6):2189–2199, November 2012.

[24] Timo Kuosmanen and Andrew Johnson. Modeling joint production of multiple outputs in StoNED: Directional distance function approach. *European Journal of Operational Research*, 262(2):792–801, October 2017.

[25] Timo Kuosmanen and Andrew L. Johnson. Data envelopment analysis as nonparametric least-squares regression. *Operations Research*, 58(1):149–160, 2010.

[26] Timo Kuosmanen, Antti Saastamoinen, and Timo Sipiläinen. What is the best practice for benchmark regulation of electricity distribution? Comparison of DEA, SFA and StoNED methods. *Energy Policy*, 61:740–750, October 2013.

[27] Timo Kuosmanen, Antti Saastamoinen, and Timo Sipiläinen. What is the best practice for benchmark regulation of electricity distribution? Comparison of DEA, SFA and StoNED methods. *Energy Policy*, 61:740–750, October 2013.

[28] Chia-Yen Lee, Andrew Johnson, Erick Moreno-Centeno, and Timo Kuosmanen. A More Efficient Algorithm for Convex Nonparametric Least Squares. *SSRN Electronic Journal*, 2011.

[29] Kamil Makiela and Jacek Osiewalski. Cost Efficiency Analysis of Electricity Distribution. *The Energy Journal*, 0(Number 4), 2018.

[30] Jacek Osiewalski and Mark Steel. Robust bayesian inference in l q -spherical models. *Biometrika*, 80, 06 1993.

[31] Rodney E. Stevenson. Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics*, 13(1):57–66, 1980.

[32] The European Commission. *Third Energy Package*. The European Commission, 2009.

Appendix A

Figures

A.1 Introduction

Term	Definition
Cost plus regulation	Cost plus based regulatory approach focuses on the realised costs, which are passed through into the allowed regulatory income or tariffs.
Direct incentives	The NRA explicitly incentivizes specific behaviour or technology, e.g. by granting a higher rate of return for certain assets or technologies.
Indirect incentives	The NRA implicitly incentivizes a certain behaviour by incentivising overarching regulatory goals, e.g. efficiency incentives may lead to more innovation together with cost reductions.
Input-based regulation	Input-based regulatory approach focusses on costs or processes, where the NRA may prescribe how certain investments are done or projects conducted.
Output-based regulation	Output-based regulatory approach focuses on parameters that describe the distribution task of the DSO or focuses on the performance of the DSO for achieving any regulatory aim. The NRA may set thresholds for relevant parameters to incentivise the DSO in a certain direction.
Price-/revenue cap regulation	The NRA ex-ante determines a regulatory allowance (price or revenue cap) for the DSO which forms the basis for the DSO's allowed revenues recovered through the tariffs charged on third parties for using its network infrastructure. The regulatory allowance is based on the DSOs individual cost structure, considering cost efficiency targets. With the calibration of the cap regulation, regulatory objectices (e.g. in terms of efficiency, quality of supply or innovation) may be calibrated.
Standardised cost regulation	Efficient costs are defined through engeneering experience. This could for example be done by calculating the involved (efficient) costs of the existing of required grid.
Technologically neutral regulation	Regulatory incentives do not create any bias towards a certain technology or cost category (e.g. CAPEX vs. OPEX).
Totex regulation	Allowed revenues do not differentiate between CAPEX and OPEX, bust considers the whole costs, instead. Therefore, it ensures that the incentive is technologically neutral.
Whole System Approach	Approach that focuses on the "system" concept, trying to identify the net benefit that regulatory decisions may bring for the whole electricity system
Yardstick competition	See cap regulation. Here, the regulatory allowance is based in parts or in total on exogenous (efficient) cost structures, for example of other DSOs.

Figure A.1: Regulatory Regimes

A.2 Literature review

A.3 Empirical results

A.3.1 DEA

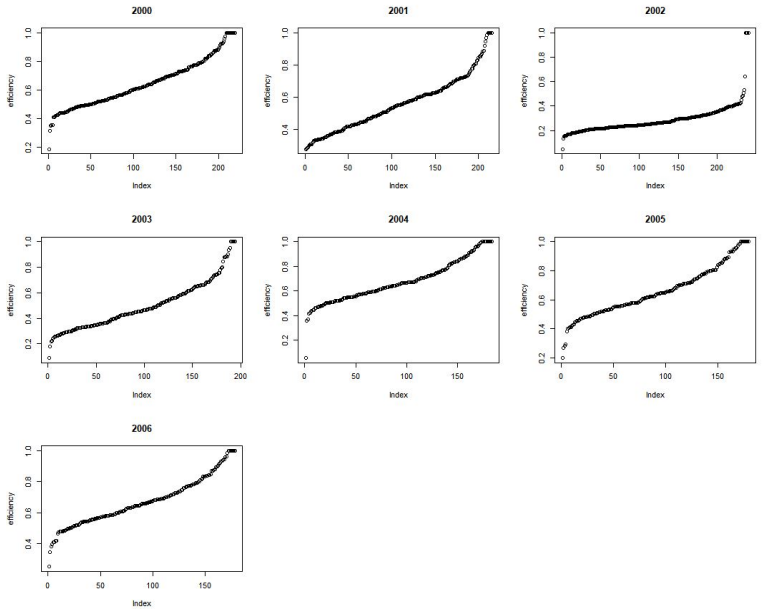

Figure A.2: Sorted Efficiencies Per Year, constant return to scale DEA

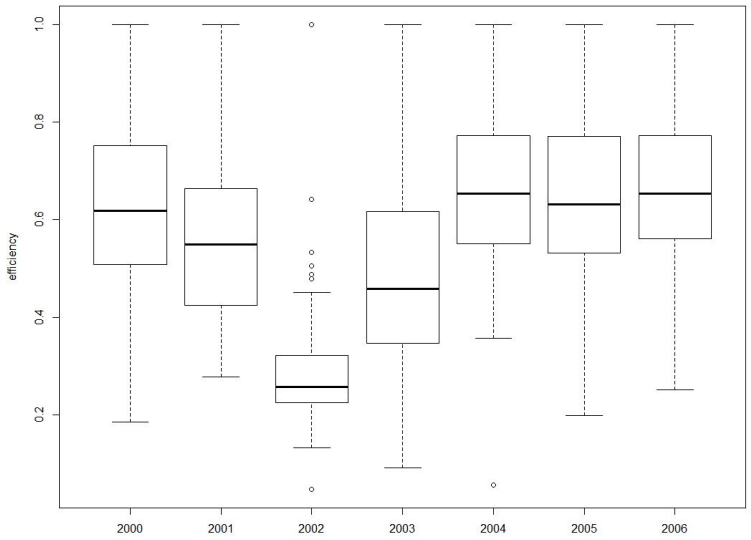

Figure A.3: Boxplots of Efficiencies, constant return to scale DEA

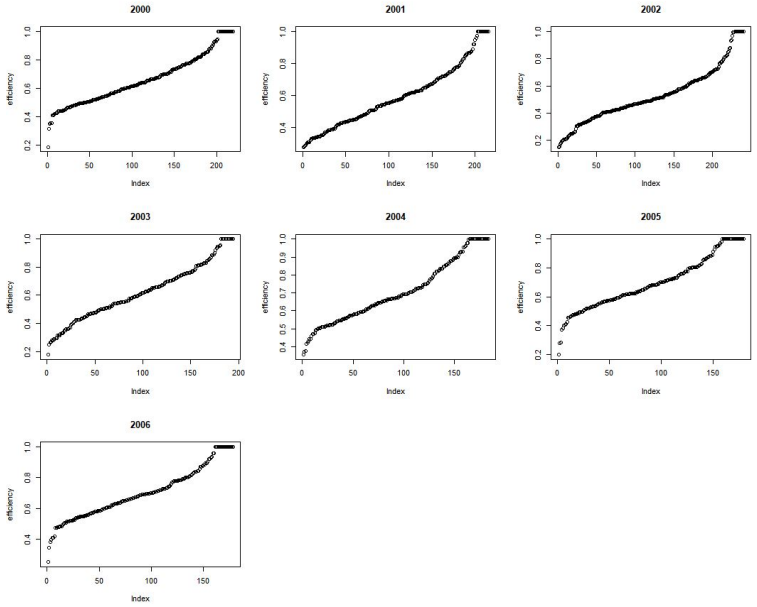

Figure A.4: Efficiencies of decreasing return to scale DEA

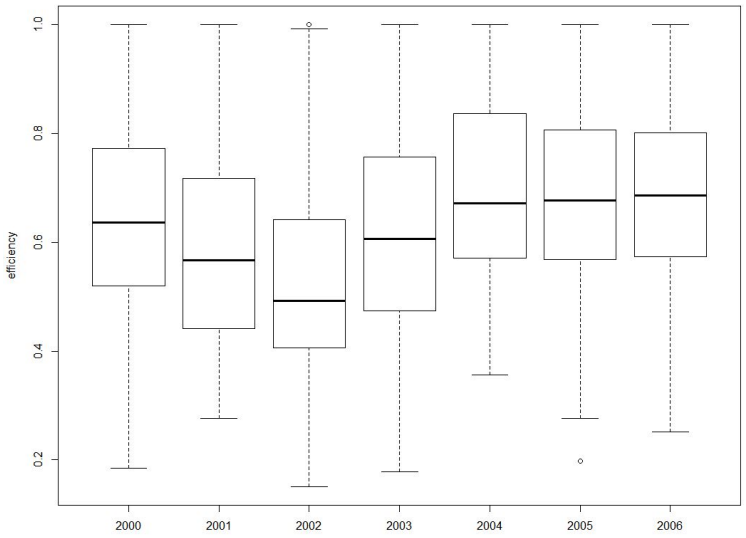

Figure A.5: Boxplots of Efficiencies, decreasing return to scale DEA

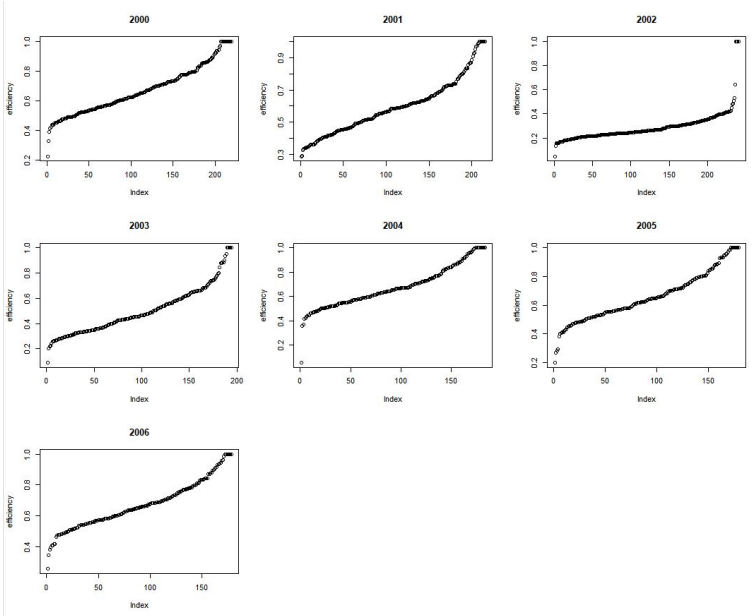

Figure A.6: Efficiencies of increasing return to scale DEA

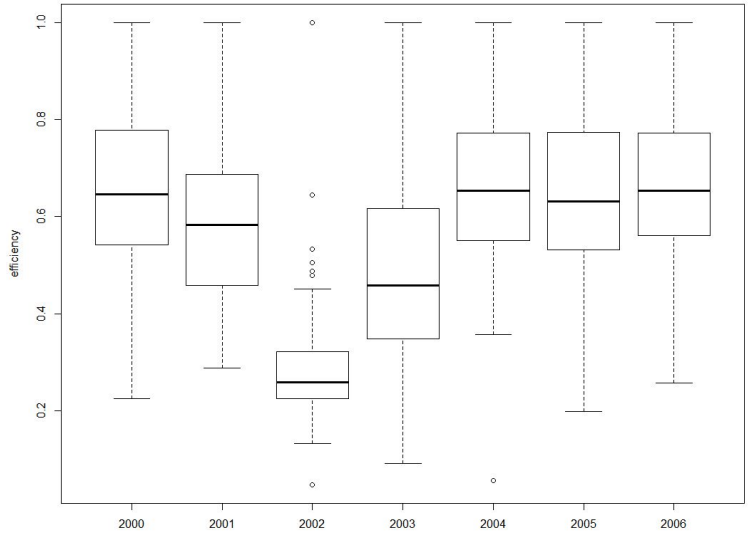

Figure A.7: Boxplots of Efficiencies, increasing return to scale DEA

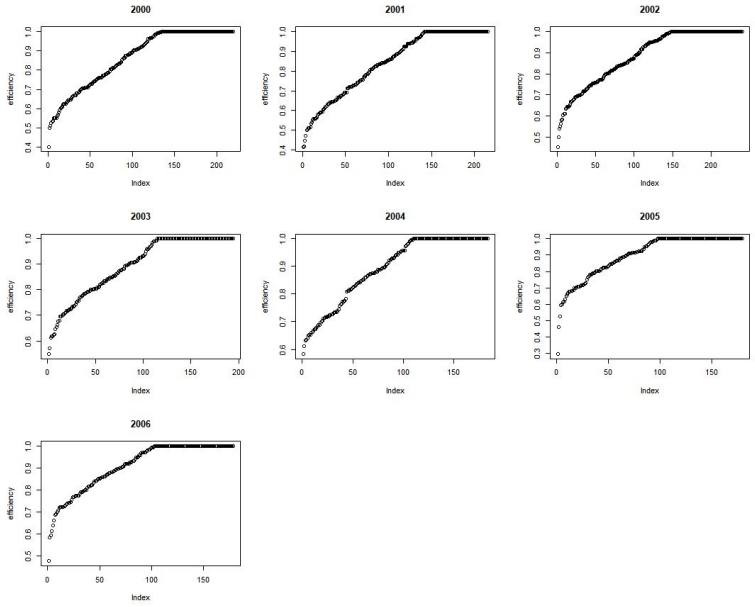

Figure A.8: Efficiencies of free disposability hull return to scale DEA

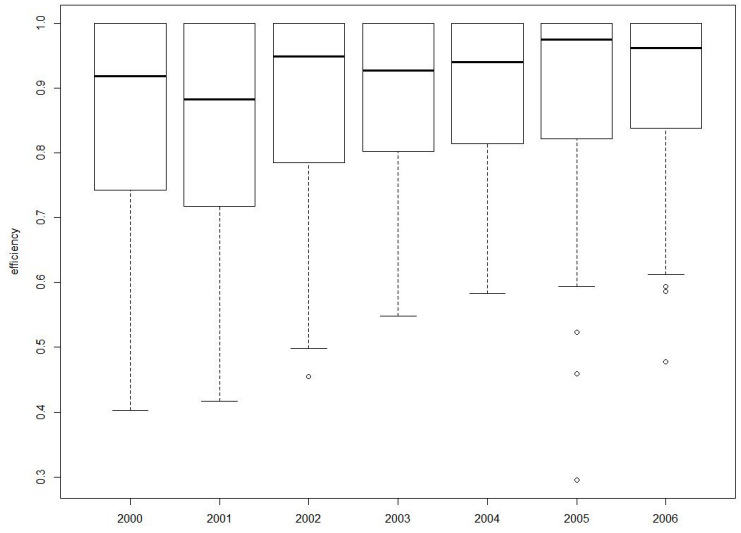

Figure A.9: Boxplots of Efficiencies, free disposability hull return to scale DEA

A.3.2 SFA

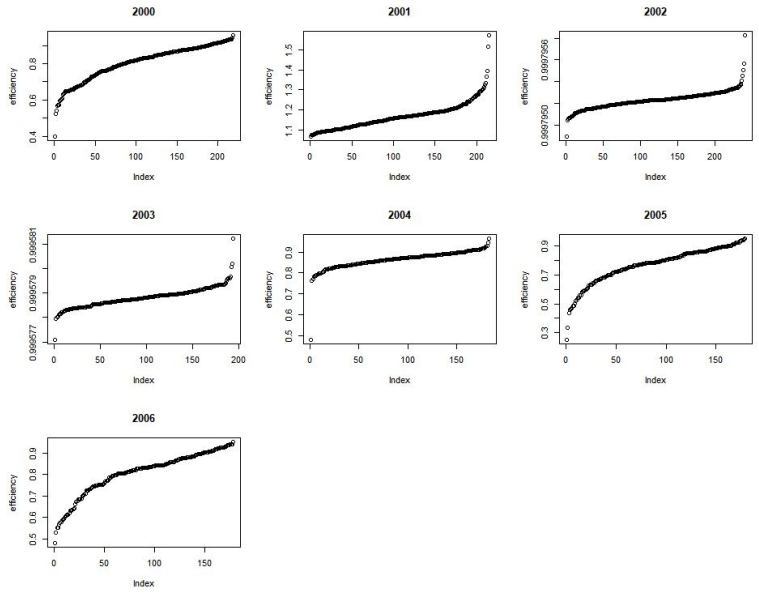

Figure A.10: Sorted efficiency for SFA input function

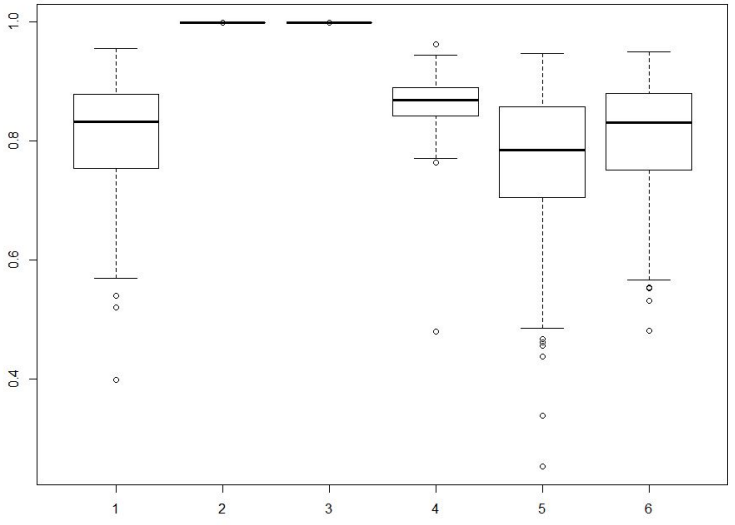

Figure A.11: Boxplot of efficiencies for SFA input function

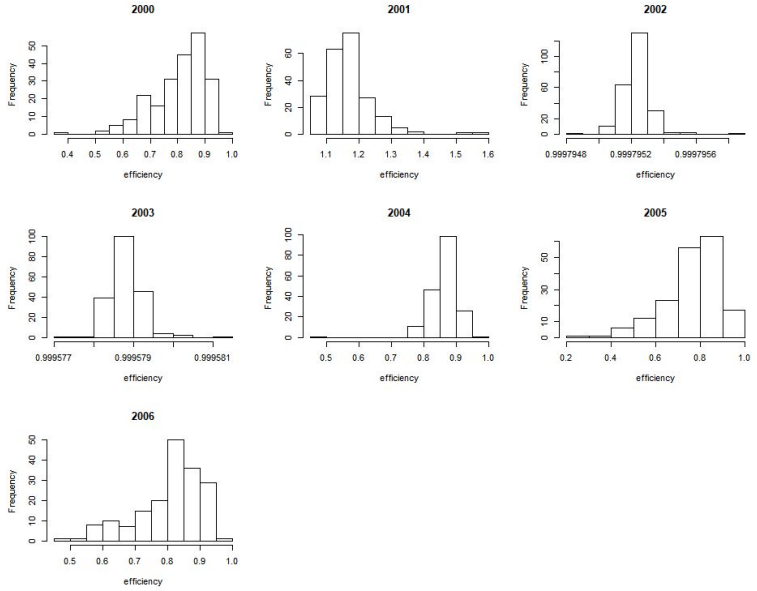

Figure A.12: Histogram of efficiency for SFA input function

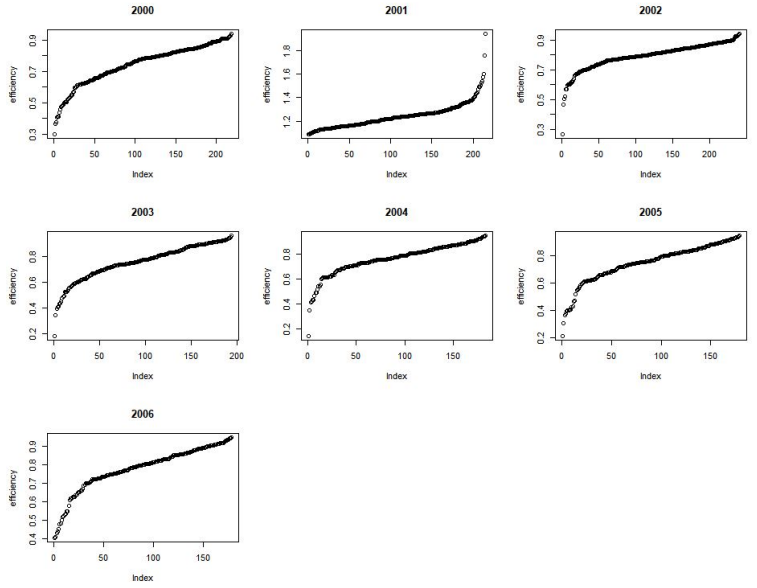

Figure A.13: Sorted efficiency for SFA output function

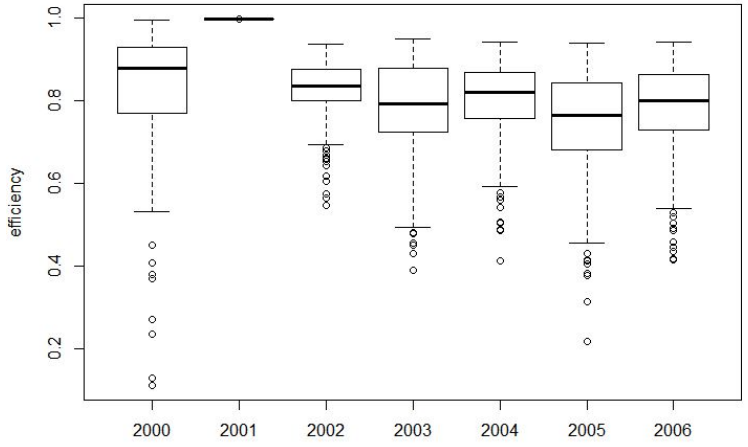

Figure A.14: Boxplot of efficiencies for SFA output function

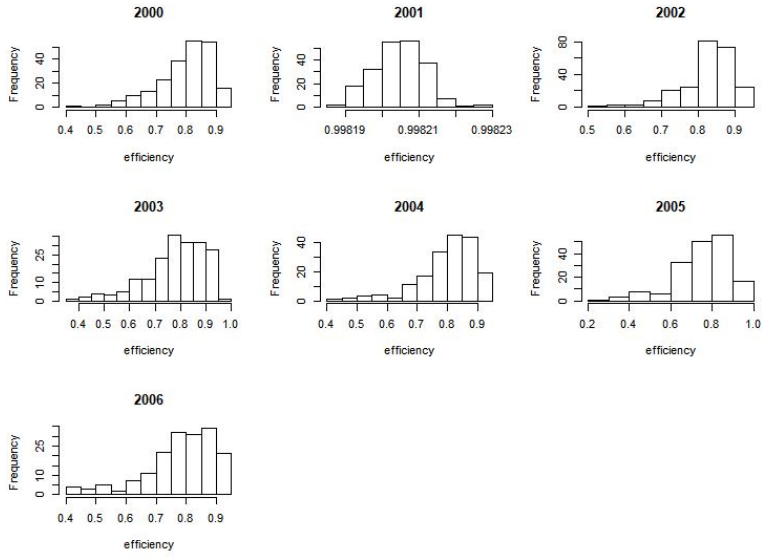

Figure A.15: Histogram of efficiency for SFA output function

A.3.3 StoNED

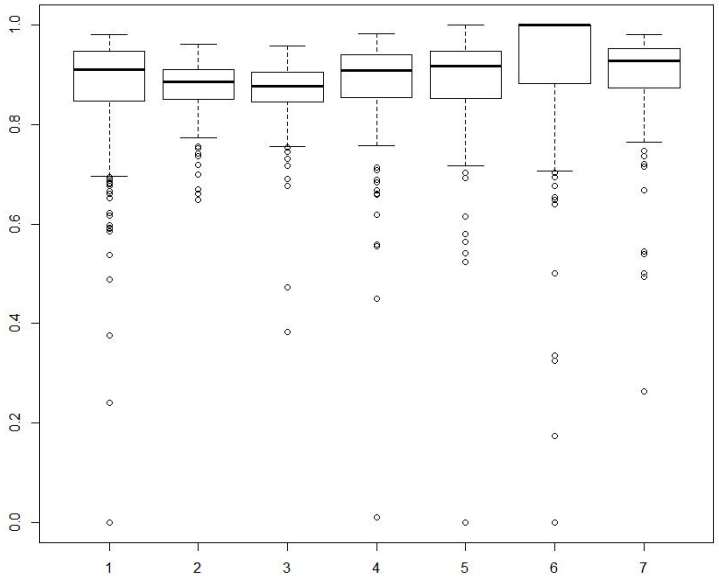

Figure A.16: Boxplot of efficiencies for StoNED model

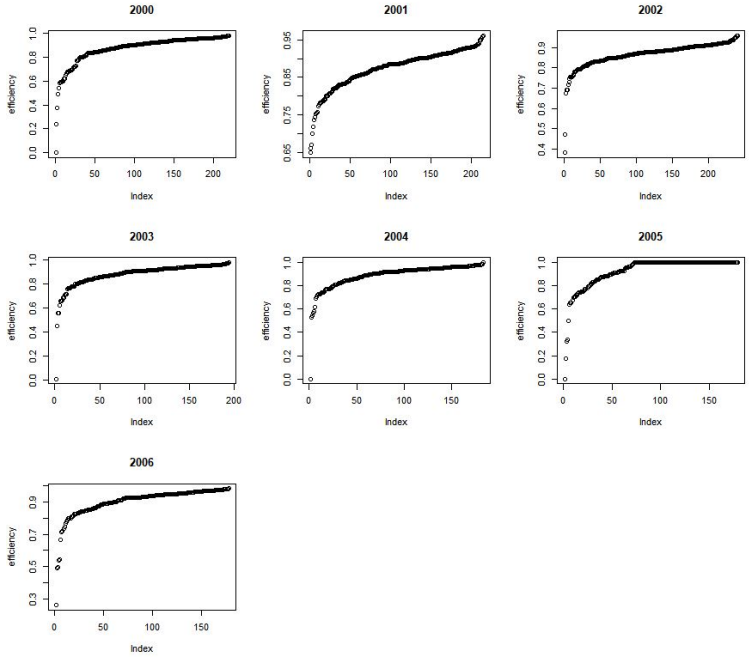

Figure A.17: Sorted efficiencies of StoNED model

A.4 Analysis

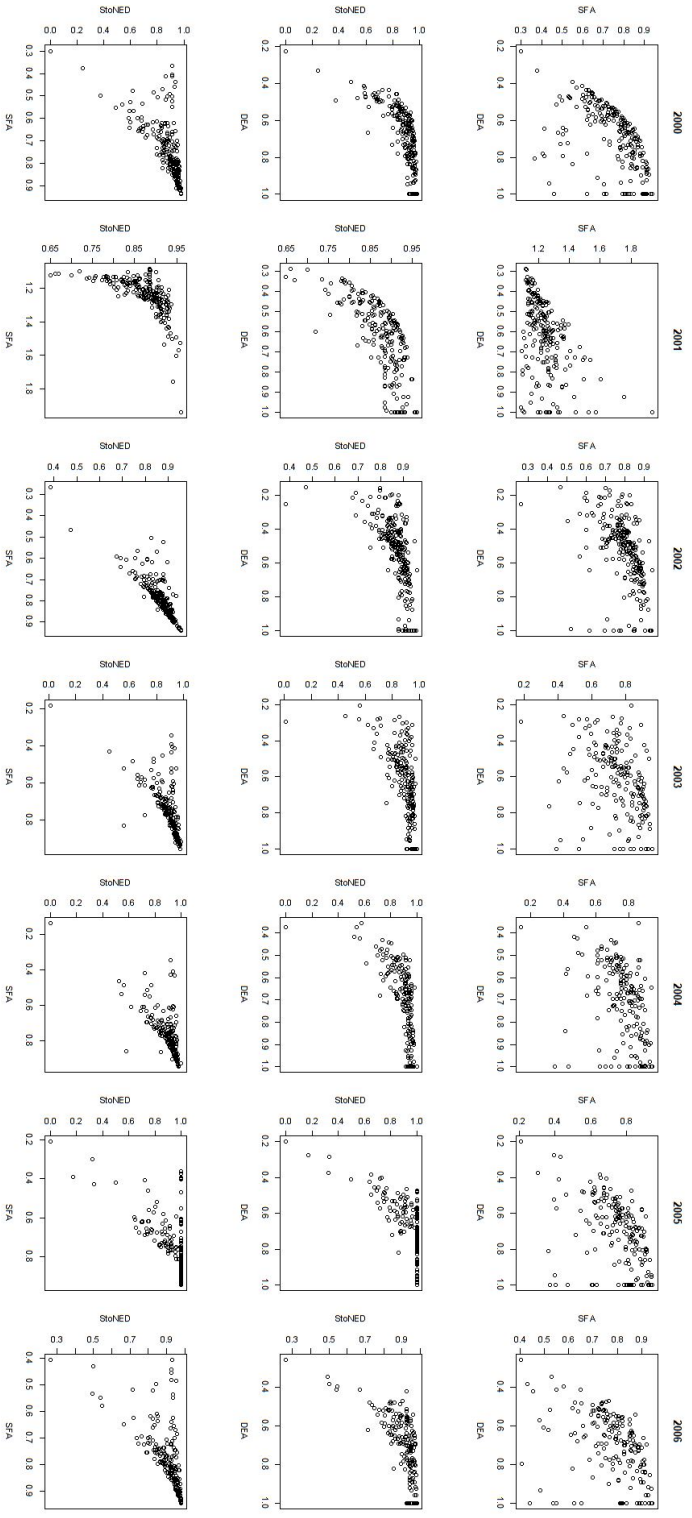

Figure A.18: Plotted efficiencies of the different models

A.4.1 Code

```
n = 50
mean_ = 1000
sd = 100

inp = cbind(rnorm(n,mean_,sd),rnorm(n,mean_,sd),rnorm(n,mean_,sd))
s = cbind(rnorm(n,mean_,sd))
s_mat = matrix(s,ncol = n,nrow = n)

res_unsort = tryCatch(stoned_dso(inp,s_mat))

inp = cbind(sort(inp[,1]),sort(inp[,2]),sort(inp[,3]))
s = sort(s)
s_mat = matrix(s,ncol = n,nrow = n)

res_sort = tryCatch(stoned_dso(inp,s_mat))
```

Figure A.19: Comparison of unsorted and sorted data

```
> res_unsort$R_sq
[1] 0.1089806
> res_sort$R_sq
[1] 0.975912
```

Figure A.20: Results of comparison of unsorted and sorted data

LOUVAIN-LA-NEUVE | BRUXELLES | MONS | TOURNAI | CHARLEROI | NAMUR
Place des Doyens, 1 bte L2.01.01, 1348 Louvain-la-Neuve, Belgique | **www.uclouvain.be/lsm**