Paper 3405-2015

Prioritizing Feeders for Investments: Performance Analysis Using "Data Envelopment Analysis" - DEA

Victor Henrique Duarte de Oliveira, Iguatinan Gischewski Monteiro, Companhia Energética de Minas Gerais – Cemig Distribuição.

ABSTRACT

This paper presents a methodology developed to define and prioritize feeders with the least satisfactory performances for continuity of energy supply, in order to obtain an efficiency ranking to support decision-making process regarding the investments to be implemented. The Data Envelopment Analysis (DEA) was the basis for the development, in which was adopted the input oriented model with variable returns to scale. To perform the analysis of the feeders, data from the utility GIS and from the interruption control system was export to SAS® Enterprise Guide, where its manipulation was made possible. Different continuity variables and physical-electrical parameters were consolidated for each feeder for the years of 2011 to 2013. They were separated accordingly to the geographical regions of the concession area, accordingly to their location (urban or rural) and then grouped by the physical similarity. Results showed that 56.9% of the feeders can be considered as efficient based on the continuity of the service. Furthermore, the results allow us to identify the assets with the most critical performance and their benchmarks, and to define preliminary goals to reach the efficiency.

INTRODUCTION

The quality of electrical energy supply has received a much greater attention by costumers, regulators and, consequently, utilities in the last years. According to the Distribution Procedures, defined by the "National Agency of Electrical Energy" (ANEEL), agency that regulates the performance of distribution systems in Brazil, the quality of energy is defined by two main aspects: product quality, related to the voltage conformity; and the service quality, related to the continuity of supply. This last one represents the capacity to meet the demand continuously, with minimal number of outages. Through the view of most costumers, that's the main aspect to meet the expected performance.

The continuity is evaluated through two essential parameters: duration and frequency of outages. From these, individual indicators – DIC (Individual Outage Duration per Costumer), FIC (Individual Outage Frequency per Costumer) and DMIC (Individual Maximum Outage Duration per Costumer) – and collective indicators – DEC (Equivalent Outage Duration) and FEC (Equivalent Outage Frequency) – are defined. Once the indicators are calculated, the Agency is capable to compare the performance of the utilities and define new goals.

In order to ensure the improvement of power supply, these goals are determined on an annual reduction, requiring the utilities to have a better operational acting, which can be reached through maintenance and investment improvements on the distribution assets. The maintenance and investment plans are main aspects in the asset management of energy utilities, that impact in their performance. These plans should be conducted in an integrated manner and are critical since they require a considerable amount of financial resources.

The planning becomes much more laborious when the volume and diversity of assets reaches greater dimensions. The Cemig D serves near 8 million customers units through its, approximately, 500 thousands kilometers of distribution network (the largest in Brazil) divided in 1686 power feeders. This network is segregated into seven main regional networks, covering 567.478 km² of Minas Gerais territory, as shown in Figure 1.

Figure 1 - Cemig D concession area

Therefore, this paper aims to present a methodology developed to support the targeting decisions regarding maintenance and investments plans in Cemig D power feeders. The methodology is based in Data Envelopment Analysis – DEA –, a Benchmarking method used to determine the efficiency of the utility power feeders, through its performance from the point of view of continuity of power supply, also considering physical and structural aspects. Thus, it becomes possible to define those who require greater investments to meet regulatory requirements and expected performance levels by the customers.

METHODOLOGY OVERVIEW

Figure 2 presents an overview of the methodology in a flowchart format to assist the understanding of each stage. These seven stages represent a manner to organize the analytic reasoning in order the help the definition of the maintenance and investments plans.

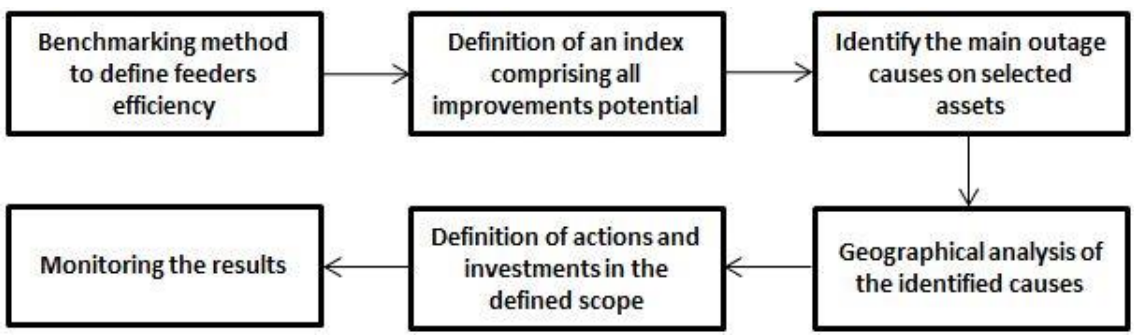

Figure 2 - Flowchart of the proposed methodology around assets management

In the following sessions, a brief description of each step of the flowchart is presented with a greater detail about the first stage, related to the benchmarking method that represents the biggest difference from the traditional process of the company in the preparation of such plans. Furthermore, the results from the first step of the methodology will be presented.

BENCHMARKING METHOD

The two main Benchmark methods were identified during the methodology development: DEA and SFA – Stochastic Frontier Analysis. Both aim to estimate quantitatively the efficiency of a group of DMUs (Decision Making Units), responsible for a process that uses a set of inputs to produce a set of outputs (products). This concept is usually extrapolated and these methods are commonly used in processes

where there is no actual physical transformation of the inputs in the products. The main difference highlighted by most authors is that SFA is a parametric model and require a hypothetical production frontier, which may or may not be close to the truth, while the DEA is a nonparametric method and no assumptions have to be made on the specific form of the border. In general, when the assumptions made are met, parametric methods have better performance. Otherwise the nonparametric methods are a better choice. Considering the complexity of defining the distribution process, DEA was chosen as the Benchmarking method.

Initially presented by (Charnes et al. 1978), the first DEA model was known as CRS, since it admitted Constant Returns to Scale. Subsequent works considered alternative hypothesis, especially (Banker et al. 1984), who proposed a method for Variable Returns to Scale – VRS. DEA consists of linear programming based method, which objective function is the maximum production given a fixed amount of inputs (product oriented model) or a minimum amount of inputs for a desired quantity of products (input oriented model). The implementation of DEA in SAS is simple and was done similarly to (Sadiq 2011).

Since it doesn't require the definition of prices or weights for inputs and outputs, as well as the knowledge regarding the explicit relation between the variables, DEA has been applied in several areas, including the power sector (Pessanha et al. 2007) - especially in regulation (Bogetoft, Otto 2011) - professional education (Lorenzett et al. 2004), military (Roll et al. 1989) and health sector (Chirikos, Sear 2000). In the Brazilian power sector, DEA was first used by ANEEL to determine the efficient operational expenditures of the utilities in the third tariff review cycle (ANEEL 2011). In that case, it was used an input oriented model and non-descending returns to scale (NDRS), considering the network extension, the number of costumers and the market as product variables, and the amount of operational expenditures as inputs.

Besides the presented advantages, DEA allows us to identify the production units with the best practices in the business, benchmarks, and to determine goals to reach the efficiency frontier. In a second stage of the methodology, these goals can be used to determine the potential improvements of each input or to construct a single metric to be evaluated, for example, a financial index comprising all input variables.

Input and Output Selection

In order to analyze the efficiency of Cemig D power feeders, twelve variables were considered relevant to measure their performance. The data used correspond to mean values of the period from January 2011 to December 2013. The mean was used aiming to mitigate the effect of severe isolated occurrences that may have happened in the time span that could compromise the analysis, like the unusual storm that occurred in June 2011. The average value is generally used when dealing with quality variables. The use of historical data despite assumptions or knowledge (that may or may not be exactly the truth) provides the robustness of this method. Having access to the data and the proposed set of analysis may grant a more assertive and efficient definition of the maintenance and investment plans by the decision makers.

In a second moment, using the DEA-Stepwise method, proposed by (Kitelsen 1993), to select relevant variables, that number was reduced to ten variables, as shown in Table 1, with their identification of input or output. To ensure the convergence of the programming, it was observed the condition of a minimum number of feeders, equal to the triple of the number of variables in the analysis (Cooper et al. 2000).

From the preliminary variables, the DEA-Stepwise model showed that the segregation of the network extension by its constructive type is as relevant as expected, since this characteristic has great influence on the continuity performance of a feeder. Beside the variables presented, some were tested regarding its installed power capacity, number of protective equipment and the segregation of the number of customers by the supply voltage.

Input/Output	Variable	Dimension
Input	DIC/FIC/DMIC Financial Compensation	R$
Input	FSS – Sustained simple frequency	Nº
Input	$CH_{Accidental}$ – Number of Costumers X Off time by accidental causes	Hours
Input	$CH_{Programmed}$ – Number of Costumers X Off time by programmed causes	Hours
Input	Quantity of Emergency Services	Nº
Output	Network Extension – Conventional	km
Output	Network Extension – Protected	km
Output	Network Extension – Isolated	km
Output	Network Extension – Underground	km
Output	Number of Costumers	Nº

Table 1 - DEA variables used in the analysis

Orientation and Model Selection

Taking into account the variables selected, an input oriented model was chosen since those are the attributes that the utility has manageability over and the reduction reflects the improvement needed to reach better performances index. The VRS model was adopted since it is considered to be less restrictive and for large samples it tends to converge on actual border. It was shown by (Banker 2011) that the VRS model can be more efficient than the NDRS regarding the frontier estimation when the real model presents non-descending returns to scale.

The Equation 1 presents the mathematical formulation for DEA-VRS input oriented model as known as Envelopment Model.

$$\begin{aligned}
&Min\ \theta \\
&s.a.: \\
&\sum_{j=1}^{n} \lambda_j x_{ij} \leq \theta x_{i0} \quad i = 1,2,\dots,m \\
&\sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{i0} \quad r = 1,2,\dots,s \\
&\sum_{j=1}^{n} \lambda_j = 1 \\
&\lambda_j \geq 0 \quad \forall j
\end{aligned} \tag{1}$$

Where:

θ represents the efficiency of the unit in analysis (vary between zero and one);

x_{ij} represents the i-th input of j-th unit;

y_{ij} represents the i-th output of the j-th unit; e

λ_j represents the shadow price or the percentage participation of the j-th unit as benchmark of the unit in analysis.

A feeder or unit is considered efficient when its efficiency is equal to one and if all input and output slacks are equal to zero. The slack is the amount of a variable that a unit can reduce (if input) or increase (if output) remaining in the efficiency frontier. If the slack isn't null for a unit in the frontier it is called weakly

efficient. Figure 3 illustrates a generic example of this case, where the unit A could improve the product 2 remaining in the frontier. If the efficiency is less than one, the feeder is considered inefficient.

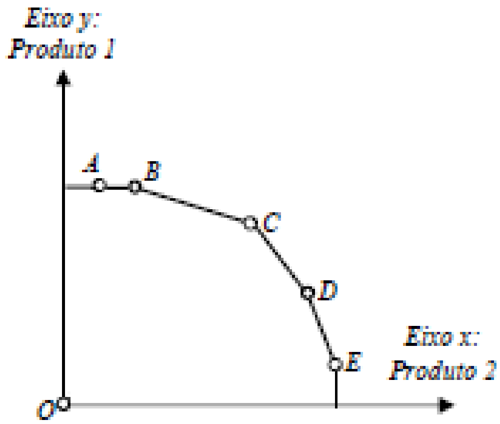

Figure 3 - Slack representation for productive unit A

Clusters of Feeders

Some of the main requirements for DEA are that the units under analysis must be comparable and must operate under the same conditions. Considering the large number of feeders and the high heterogeneity between some of them and between the areas of operation, clustering becomes necessary.

In a first step, the feeders were separated into the seven regions that compose the Cemig D concession area. The geographic division of the distribution network eases the effect of environmental variables, like precipitation and lightning, which have a large impact on the performance. If data related to such variables were available, they could be included in the model or used to rectify the efficiency results. Furthermore, it is interesting to segregate in this networks that are commonly used to consolidate investments, in other analysis made by the company.

Then, the feeders were dismembered according to their location – urban or rural. After this, the initial scenario of 1686 units was expanded to 2799 units. This division is explained by the difference between the field services in these areas, especially due to the dispersion of the network, which has a large influence on the input variables. Thus, feeders' portions are only compared with others in the same location as the reference.

At last, within the groups previously created, some clusters were defined so that the analysis is done only considering similar feeders. The clustering method used is based on the one used by ANEEL to define the goals of the continuity indices (DEC and FEC) of the electrical assemblies of the utilities (ANEEL 2012). Applying the concept of Euclidean distance for the physical-electrical variables, it was determined, for each feeder, the group of those who have the similar characteristics, limited to a maximum heterogeneity index.

Besides all this grouping, it is noteworthy that 28 predominantly underground feeders were disregarded in the analysis, since they are in other technological level compared to most of feeders. Their high performance levels would be crucial in estimating the efficiency frontier, compromising the results.

DEFINING A SINGLE IMPROVEMENT INDEX

In some DEA applications the efficiency score given for each DMU may be enough to rank the main targets for improvement actions, but, for managing purposes, the construction of a new index regarding financial values could be useful. The Equation 2 presents the definition of such index here called FPI – Financial Potential Index:

$$FPI_i = v_{1,i} + v_{2,i} * c_{1,i} + [\left(v_{3,i} + v_{4.i}\right) * c_{2,i} * c_{3,i}]/8760 \qquad (2)$$

Where:

$\mathbf{v_{1,i}}$ is the potential reduction of financial compensations in feeder $\mathbf{i}$

$\mathbf{v_{2,i}}$ is the potential reduction of emergency services in feeder $\mathbf{i}$

$\mathbf{v_{3,i}}$ is the potential reduction of $CH_{Accidental}$ in feeder $\mathbf{i}$

$\mathbf{v_{4,i}}$ is the potential reduction of $CH_{Programmed}$ in feeder $\mathbf{i}$

$\mathbf{c_{1,i}}$ is the average cost of an emergency service in feeder $\mathbf{i}$

$\mathbf{c_{2,i}}$ is the average annual market consumption in feeder $\mathbf{i}$

$\mathbf{c_{3,i}}$ is the average tariff in feeder $\mathbf{i}$

The first and second parts of the equation are intuitive. The last one represents the cost of the power that isn't supplied. In this case, the reduction in the number of outages was not included in the definition of the FPI because its potential reduction is almost equivalent to the emergency service reduction potential.

IDENTIFYING THE MAIN CAUSES AND DEFINING ACTIONS

Once the FPI is created and the best targets are determined, it is important to identify the main outage causes and their effects on the index of the selected feeders. This analysis serves to identify offenders that cause the greatest influence on the quality of service. With the information presented, the analyst may direct the planning according to the reduction potential of each group of causes.

It should be noted that the expert analysis is indispensable to ensure the possibility of acting on the causes. For example, a feeder with a FPI well distributed in various types of causes can be more difficult to be improved if compared to a feeder with smaller FPI concentrated in small number of types of causes.

After the cause analysis, it is necessary to check if there are any geographical factors that could be concentrating outages in a single area. This kind of analysis is the action planning since it can provide a systemic solution instead of an individual on, covering more feeders with fewer resources.

MONITORING THE RESULTS

Once the maintenance and investment plans are defined and executed, it is necessary to monitor the results. The proposal is to use this methodology dynamically, by having the data set updated constantly, allowing the time window shift so that the results can be compared and the improvements to be identified for each planned action. Furthermore, the continuous analysis is essential to allow the correction of the implementation of planning, in a timely manner, if any deviation in the expected results is identified.

The implementation of this methodology in a SAS® Enterprise Guide project allows the automatic update in the results once the data exported by other corporate systems is being routinely updated in the libraries.

BENCHMARKING METHOD RESULTS

The results of the analysis using the DEA method showed that from the 2799 partial feeders considered, 1592 (56,9%) were efficient while 1207 (43,1%) were inefficient regarding the continuity of power supply. The mean efficiency score for the feeders was 0,745. The Table 2 summarizes the results for each one of the regions, showing the number of feeders identified in each efficiency classification.

Region	Efficient	Inefficient
Center	321 (60.0%)	214 (40.0%)
North	216 (59.3%)	148 (40.7%)
South	225 (57.3%)	168 (42.7%)
East	227 (51.5%)	214 (48.5%)
West	202 (59.1%)	140 (40.9%)
Mantiqueira	170 (61.4%)	107 (38.6%)
Triangle	231 (51.7%)	216 (48.3%)

Table 2 - Benchmark results for the network regions

The Figure 4 complements Table 2, presenting an efficiency level histogram for all feeders consolidated. It is clear that some few feeders with extremely low efficiencies might be the main targets of decision planning.

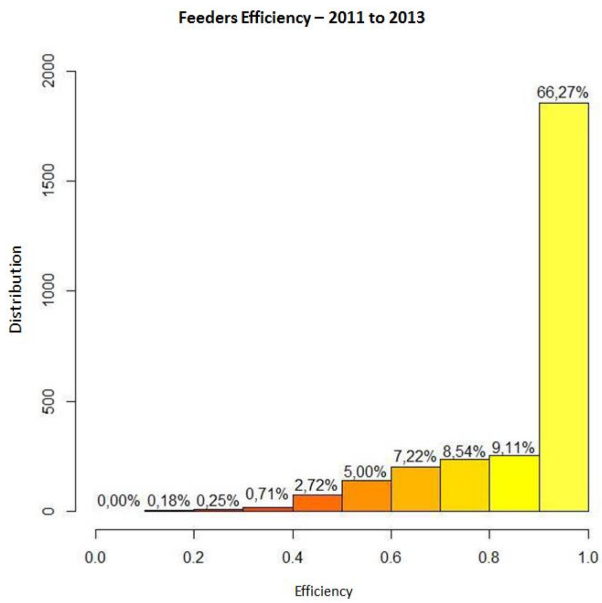

Figure 4 - Efficiency histogram for all partial feeders. Data from 2011 to 2013

Furthermore, as mentioned, the methodology allows an individual analysis of the feeders with lowest efficiency levels, through identification of benchmarks and input reduction potential. Table 3 exemplifies these observations.

	Financial Compensation (R$)		FSS		CH Accidental		CH Programmed		Emergency Services	
Feeder	Current	Goal	Current	Goal	Current	Goal	Current	Goal	Current	Goal
GVSD10 (R)	9162,51	1265,89	332	103	91451,08	8621,9	10816,55	3636,2	920	309
CRL 12 (R)	4050,45	1471,01	276	100	47203,74	13376,41	12860,59	3766,065	411	149

Table 3 - Inputs improvement potential and benchmarks identification

Once defined the goals for each one of the inputs, it is also possible to estimate the impact of them over the key performance indicators, if reached. For example, the difference between the goal and the current situation for the “CH” variables allows us to evaluate the effect over the DEC indicator. Analogously, it is possible to estimate the expected reduction of the operational expenditures once the emergency service goal is reached. These inferences are important to define the possible improvement solution for each feeder in order to measure the real possible gains in each scenario.

CONCLUSION

The definition of assets maintenance and investment plans is essential for the full exercise of an utility activity, allowing the service to be performed with excellence to its customers. This paper presented a methodology based in Data Envelopment Analysis – DEA – to determine the efficiency levels of power feeders, as well as define a preliminary ranking to support the decision making regarding an investment implementation in the distribution assets. The proposed methodology allows the planning to be carried out more dynamically compared to what is commonly practiced.

To build the model, we used data of the distribution’s physical structure and the historical data of power outage. All these information were exported from the utility systems to SAS® Enterprise Guide, which enabled the data manipulation to consolidate all desirable variables for each feeder.

The definition of the performance of each feeder is essential to the asset management and the DEA proved to be very effective in order to identify the ineffective units. Once the benchmarking analysis proposed in the methodology is made, the teams responsible for investment, operation and maintenance, can make their specific analysis, such as root cause and geospatial analysis, for those identified feeders, so that, finally, the more consistent and grounded decisions can be made. In addition, the input goals resulted by the method can be used as estimative to help the definition of the amount of investment to be used in each case. The proposed methodology is expected to be able not only to meet the specific demands for planning and maintenance works but to spread the culture of analytics within the company.

This analysis is currently being used in parallel with the traditional methodology used to define the maintenance plans in order to support and confirm the best chosen actions. In future works, this methodology should be used to analyze others performance parameters – such as voltage quality, power

losses or commercial service – aiming to comprehend other aspects and important criteria in the definition of greater investment plans.

REFERENCES

National Agency of Electrical Energy – ANEEL. October/2011. Technical Note nº 294/2011 – SRE/ANEEL. Public Hearing 40/2010.

National Agency of Electrical Energy – ANEEL. December/2012. Technical Note nº 193/2012 – SRD/ANEEL.

Banker, R.D. 2011. Report on ANEEL's Proposal for Electricity Distribution Tariff Regulation Technical Note No. 101/2011.

Banker, R.D.; Charnes, A. & Cooper, W.W. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science, 30: 1078-1092.

Bogetoft, P & OTTO, L.(2011). Benchmarking with DEA, SFA and R. International Series in Operations Research & Management Science 157.

Charnes, A.; Cooper, W.W. & Rhodes, E. 1978. Measuring the efficiency of decision making units. European Journal of Operational Research, 2: 429-444.

Chirikos, T.N., Sear, A.M. 2000. Measuring hospital efficiency: a comparison of two approaches. Health Services Research.

Cooper, W.W.; Seiford, L.M. & Tone, K. 2000. Data Envelopment Analysis, A comprehensive Text with Models, Applications, Reference and DEA-Solver Software. Kluwer Academic Publishers, Boston.

Kitelsen, S.A.C. 1993. Stepwise DEA; Choosing variables for measuring technical efficiency in Norwegian electricity distribution.

Lorenzett, J.R.; Lopes, A.L.M & Lima, M.V.A. 2004. Aplicação de Análise Envoltória de Dados (DEA) na Avaliação de Desempenho de Unidades Produtivas para Área de Educação Profissional.

Pessanha, J.F.M.; Souza, R.C. & Laurencel L.C. 2007. Um modelo de análise envoltória de dados para o estabelecimento de metas de continuidade do fornecimento de energia elétrica. Pesquisa Operacional, v.27, n.1: 51-83.

Roll, Y.; Golany, B. & Seroussy, D. 1989. Measuring the efficiency of maintenance units in the Israeli Air Force. European Journal of Operational Research, 43: 136-142.

Sadiq, S. 2011. The Final Frontier: A SAS® Approach to Data Envelopment Analysis. SAS Global Forum 2011, Las Vegas, NV.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Victor Henrique Duarte de Oliveira
Companhia Energética de Minas Gerais – Cemig Distribuição
+55 (31) 3506-2115
victor.holiveira@cemig.com.br