CLUSTER ANALYSIS AS A TOOL TO ASSESS THE PUBLIC PERCEPTION OF SOCIAL AND ENVIRONMENTAL IMPACTS OF HYDROPOWER PROJECTS[1]

Guilherme Rodrigues Lima

Carlos Eduardo Frickmann Young

Instituto de Economia

Universidade Federal do Rio de Janeiro

Email: Young@ie.ufrj.br

Abstract

This study presents the results of an ex-post assessment of economic, social and environmental impacts of the construction of the Itapebi hydropower station, in Brazil. The methodology was based on the application of cluster analysis, a statistical tool to organize data in groups of similar characteristics, to identify the different perceptions about the consequences of the dam, based on interviews carried out among the affected population seven years after its conclusion. The main objective was to identify how differently the affected people in the sample perceived the long term impacts of the project, according to geographical and socioeconomic characteristics. Indeed, the results show that interviewees presented a significantly differentiation in their answers: the ones who identified more negative aspects were those who live upstream the dam, or had their economic activities damaged by it. The main conclusion is that, during the ex-ante impact assessment required for the environmental licensing, cluster analysis based on interviews of the many diverse groups affected could be a very helpful tool to identify how differently the long term impacts are perceived, and more robust and less conflictive mitigation proposals can improve the social acceptance of the project.

Keywords: environmental impacts, hydropower, cluster analysis, Itapebi

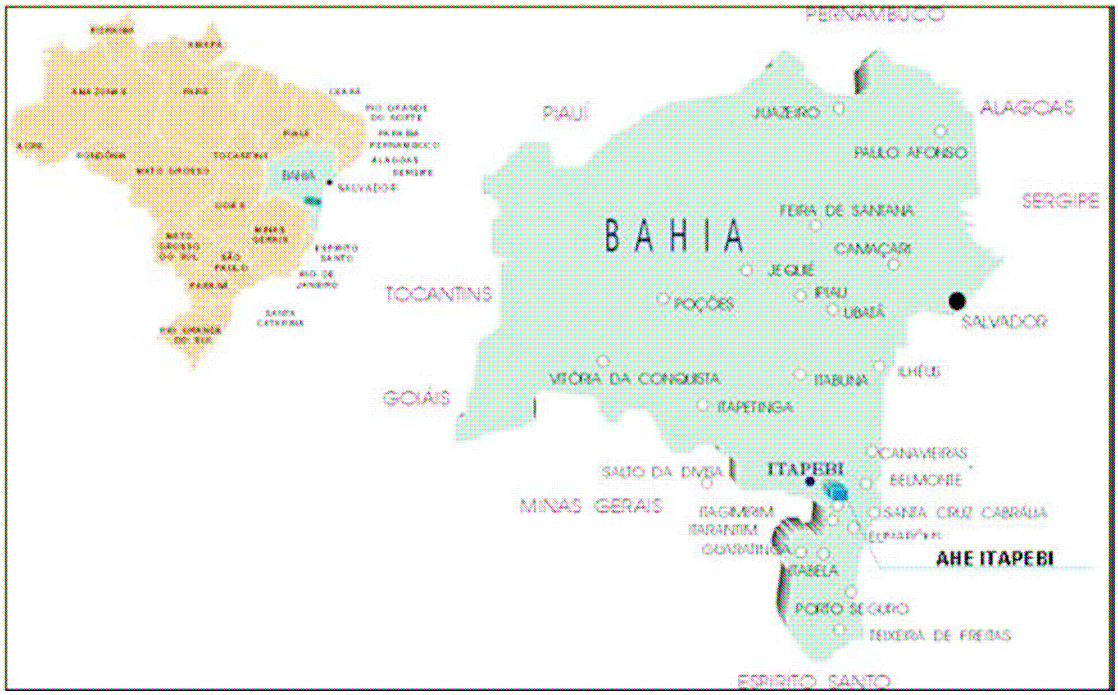
---

[1] This paper is based on the monograph elaborated by the first author (Lima 2012), supervised by the second author. We are thankful for the support provided by CBEM and Itapebi/Neoenergia.

Introduction

This study presents the results of an ex-post assessment of the public perception about the economic, social and environmental impacts of the construction of the Itapebi hydroelectric power station (Itapebi HPS) in the Northeast of Brazil (Map 1) using cluster analysis.

Map 1. Location of the Itapebi Hydroelectric Power Station



Source: Engevix *apud* Gavião (2006).

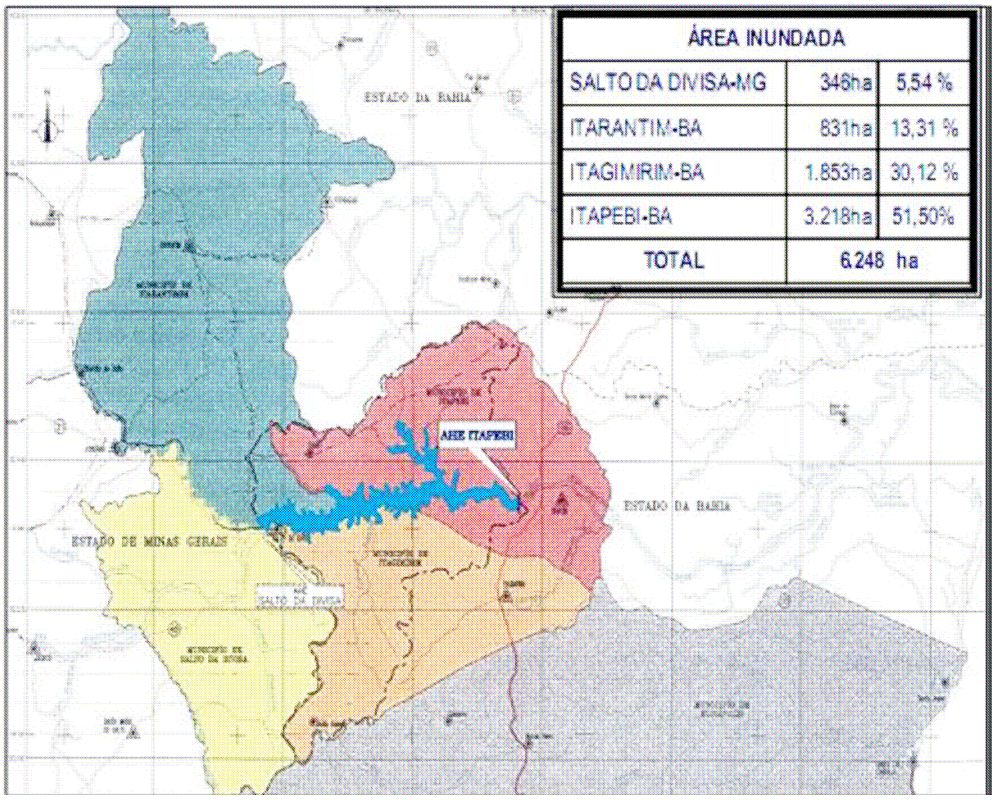
Many conflicts erupted during the construction of the Itapebi HPS during the 1997-2002 period, most of them concerning environmental impacts of the project that were not correctly anticipated by the environmental impact statement presented in order to obtain the legal authorization for the reservoir (Gavião 2006, Pereira 2012). Many years after the completion of the project, a group of researchers visited the affected region, carrying out an ex-post assessment of the impacts of the project and contrasting them to the original expectations presented in the project documentation (CBEM 2011). Among other activities, a survey was conducted in 2010 in the affected region, asking the local inhabitants about their perception of the consequences of the Itapebi HPS.

This paper presents the results of the application of cluster analysis, a statistical tool to organize data in groups of similar characteristics, to identify the different perceptions about the consequences of the dam, based on interviews carried out among the affected population seven years after its conclusion (Lima 2012). The main objective was to identify how differently the affected people in the sample perceived the long term impacts of the project, according to geographical and socioeconomic characteristics. The implicit hypothesis is that the use of cluster analysis in the assessment of the perception of the population affected by a large infrastructure project, such as a hydropower dam, can be a helpful instrument to improve the capacity of predicting the impacts of such enterprise since it has the property of elucidating how differently the many diverse social groups answer to the same questions. Understanding the diversity of the perceptions of the affected communities about the impacts of the projects can improve the design of mitigation and compensation measures.

The Itapebi Hydroelectrical Power Station

The Itapebi HPS is located in the Jequitinhonha River, in the South part of the Brazilian State of Bahia, near the border with the State of Minas Gerais. The Itapebi HPS is considered a medium size power station, with installed capacity of 450 MW, and three generators of 150 MW each (Gavião, 2006). The reservoir inundated an area of 6,248 hectares and a volume of 1.6 billion cubic meters. The Directly Affected Area (ADA) of the enterprise comprises the municipalities of Itapebi, Itarantim and Itagimirim in the state of Bahia, and Salto da Divisa in Minas Gerais (Map 2).

Map 2. *Municípios* affected by the Itapebi Hydroelectric Power Station



Source: Engevix *apud* Gavião (2006).

The ADA includes land that was inundated by the reservoir, construction sites, accommodation for workers and other buildings related to the enterprise. It also includes the urban areas that suffered changes on their socioeconomic structure.

This region around the project is highly dependent on agriculture, and it is characterized by low level of development, with low income, high unemployment and poor social indicators all components of the Municipal Human Development Index (IFDH, estimated by the United Nations Development Program) for the region are below the respective state averages.

The Environmental Impact Statement (EIA), elaborated in 1995, identified thirteen socioeconomic impacts deriving from the construction and operation of the hydropower project. The main impacts identified were:

- Changes in the demography – The creation of approximately 2.200 direct jobs and 1.000 indirect jobs would induce high migration flows to the region. These

impacts would be more intensives during the construction, and then it was expected a counter flow, with people leaving the region. The mitigating measures suggested were hiring mainly local workers, monitoring the population and a public health program.

- Changes in the housing market – The impacts in this segment were caused mainly by the migration of workers to the area, putting pressure through the demand for habitation.

- Changes in the labor market – The sudden rise of work opportunities in the region, mainly for non qualified workers, would start a boom cycle, to be followed by a later reversion after the completion of the civil onstruction activities.

- Changes in the demand for social services – This impact is related to the demographic changes, and the increasing population would make pressure over the social services, such as education, health, transport and sanitation, mainly in the *município* of Itapebi.

- Increase in the access to electricity – This was considered the main positive impact of the hydropower, with expectations of rising consumption of energy. It was recommended an assessment on how to expand the electricity distribution in rural areas, mainly around neighbor cities.

- Impacts in the urban area of Salto da Divisa – This refers to the flooding that would occur in the urban area of the city, which is located by the side of the Jequitinhonha River. The area had 50 houses and 235 people would be affected. The EIA stated that this number represented a low share of the municipal population.

- Impacts on the health system – This impact is related to the increase in diseases incidence and deterioration of the public health system. The incidence of diseases tends to grow due to the migration, changes in the environment, especially during the flooding, and the deterioration of the water quality, causing digestive diseases.

Besides these impacts predicted on the EIA, as in almost all large infrastructure projects, other unpredicted effects were observed. Some of the mitigating measures were the reallocation of some families to new houses, when their previous houses were affected; paving to recover the streets; reform of squares etc. However, many other problems that are pointed out by the population as consequence of the construction of the hydropower were not taken into consideration, including the temporary rise of the male population and subsequent rise in the number of single mothers, water pollution of the water, diseases, leisure and loosing of jobs Pereira (2011). As a consequence, there were many legal disputes concerning the project, with most of the resistance against it concentrated in Salto da Divisa. After a long period of conflicts, a settlement was finally achieved and the Itapebi HPS started its operation in February 2003.

As argued by Pereira (2011), it shows that *ex ante* studies, such as the environmental impacts statement, are not capable to predict all the impacts that would result from the enterprise and, therefore, it is necessary to establish a continuous collection of data about the project during its building stage and after it becomes operational. The next section presents the cluster analysis methodology, used by Lima (2012) to analyze the results of the survey carried out in the *municípios* affected by Itapebi HPS (CBEM 2011).

Cluster Analysis

Cluster Analysis is a technique of multivariate analysis used to explore the database or to simplify it, making easier its interpretation. It is an explanatory technique that differs from the regression analysis because it does not need previous hypothesis about the data.

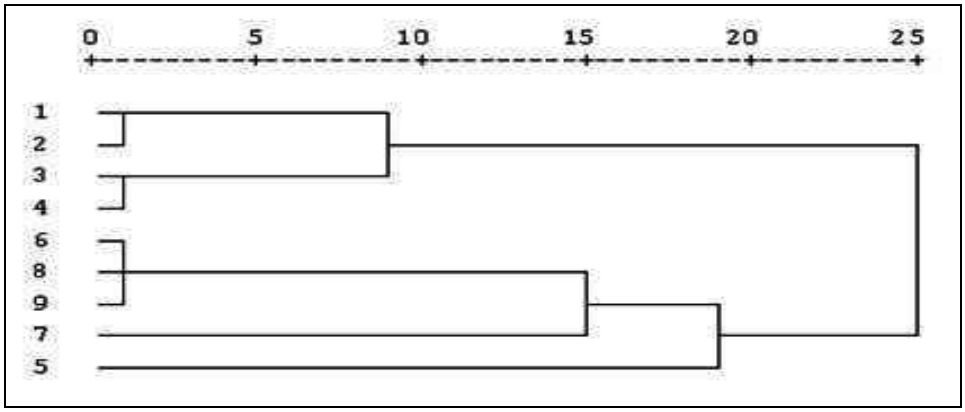
According to Mingoti (2007), the Cluster Analysis can be useful in many fields of study: in psychology, to classify people by their personalities (apud Speece at al., 1985); in market surveys; in ecology, to classify species; in geography, to classify regions, states or cities by economic, demographic and physic variables.

In this work, each person who answered the survey is characterized by the variables (questions) which were selected for the study. The posterior use of values for these

answers made it possible to establish a measure of distance between the individuals and they could then be clustered based on this measure.

The figure 1, below, exemplifies the diagram of clustering nine elements. First, the individual elements form small groups in which the components are very similar one to each other.[2] Then, new groups are formed by grouping previous clusters with single elements or with other clusters.

Figure 1. Example of a diagram of cluster analysis



Source: Lima (2012)

There are several possible methods to proceed the clustering and also different measures of distance between the elements, and explaining these techniques would run out of the scope of the paper. It is important, however, to mention that the Euclidian Distance was used as a measure of distance, and the K-means method was used to cluster the elements.

The database used in this work is composed of the answers to the surveys applied in the four cities affected by the hydropower (CBEM 2011). These surveys were part of a study whose goal was to develop ways of analyzing *ex post* socioeconomic and environmental impacts of hydropower stations. The questions aimed to identify whether the population accepted or not the measures adopted by the enterprise to mitigate or to

[2] The picture shows in a first stage three groups being formed. In fact, in each round of the clustering only two elements are grouped. This impression in the picture happens because the distance between those groups is too small.

compensate the impacts, to verify the effectiveness of such measures, as well as to valuate the changes in the quality of life of those populations.

The surveys were elaborated from the analysis of documents related to the licensing process, of impacts mentioned by the population, from the consulted bibliography and from the dissertation of Gavião (2006), who studied the conflicts between the enterprise and the population of Salto da Divisa. The surveys had two approaches: quantitative, with objective questions (closed options); and qualitative, with subjective questions, in which people had more autonomy with their answers.

There were three types of surveys: one to the general population; one to specific groups of stakeholders (people who were resettled, farmers, fishers, rock extractors, sand extractors and washerwomen); and one to public authorities. The first type addressed general aspects of the population, such as socioeconomic aspects, health, housing and also addressed their opinion about the benefits and/or damages brought by the reservoir. The second type of survey had two parts, the first with general questions about the person and the opinion about how did the reservoir affect his/her life; the second with specific questions to each group. The third type of survey, which is not analyzed here, had questions to the public authorities about the dam and its impacts on the environment, leisure, health and the financial compensations.

In this paper we used the surveys for the general population and for the specific groups. In the first case the questions (hereinafter mentioned as variables) were divided in three sets: socioeconomic; health and sanitation; and public services.

The set of socioeconomic issues includes:

i) income;

ii) job opportunities;

iii) commercial activity;

iv) unemployed people;

v) rental price;

vi) price of building materials;

vii) moves due to rise of rental prices.

The set of health and sanitation issues has four variables:

i) waste collection;

ii) quality of public health;

iii) quality of the water of Jequitinhonha River;

iv) disease incidence.

The third set, about public services, included the following variables:

i) electricity offer;

ii) public transport;

iii) quality of public schools;

iv) maintenance of streets and roads;

v) recreation areas;

vi) violence;

vii) single mothers with difficulties to raise their children;

viii) children out of school and without proper social assistance.

For the analysis of the specific groups, five questions (variables) were selected:

i) how did the dam affect your life?;

ii) how did the city was changed after the dam?;

iii) job opportunities;

iv) could you maintain the same activities you had before the dam?;

v) income.

The sample had, in total, 292 interviews for the general population divided among the municipalities as follows: 89 (30.5%) interviews in Itapebi; 19 (6.6%) in Itarantim; 58 (19.8%) in Itagimirim; and 126 (43.2%) in Salto da Divisa. The great number of surveys in the later city is explained because it was the most affected negatively by the dam, causing several conflicts between the enterprise and the population.

Concerning the specific groups, there were 72 interviews: 23 (31.9%) of people who forced to move; 15 (20.8%) of fishers; 4 (5.6%) of sand extractors; 7 (9.7%) of rock extractors; 7 (9.7%) of washerwomen; and 16 (22.2%) of farmers.

It is important to mention, however, that during the cluster analysis, the software (SPSS) removes all the elements which present missing data for any variable. For example, if one person did not respond the question about waste collection, such data will be missing in the sample and, even if the person responded all the other questions, he or she will not be considered in the analysis. For this reason, the different sets of variables (socioeconomic, health and sanitation, and public services) did not have 292 elements, as well as the numbers in each one are different. The same occurs with the specific groups.

Results

The cluster analysis identified a statistically significant pattern of answers. Using the answer to the socioeconomic questions, three groups of respondents were formed, with Group 1 being composed for those who felt mainly benefitted by the project, Group 3 composed by those who believe that were mainly prejudiced by the reservoir, and Group 2 in an intermediate position.

Table 1. Distribution of groups according to socioeconomic questions, by *município*

	G1 (benefitted)	G2 (intermediary)	G3 (prejudiced)	Total
Itagimirim	20	11	19	50
%	40,0%	22,0%	38,0%	100,0%
Itarantim	14	2	0	16
%	87,5%	12,5%	0,0%	100,0%
Itapebi	39	11	26	76
%	51,3%	14,5%	34,2%	100,0%
Salto da Divisa	17	28	63	108
%	15,7%	25,9%	58,3%	100,0%
Total	90	52	108	250
%	36,0%	20,8%	43,2%	100,0%

Source: Lima (2012)

As expected, the more negative perception of the population was observed in Salto da Divisa, located upstream the dam and were most of the environmental impacts were concentrated - not surprisingly, almost all of the resistance against the project was located in Salto da Divisa (Gavião 2006, Pereira 2011). On the other hand, the more favorable answers were concentrated in the *municípios* that receive more royalties from the reservoir, but were much less affected by the project.

Similar results were obtained when the questions referred to health and sanitation issues (Table 2), and the quality of public services (Table 3), with the worse perception declared by the inhabitants of Salto da Divisa (Lima 2012).

Table 2. Distribution of groups according to health and sanitation questions, by *município*

	G1 (benefitted)	G2 (intermediary)	G3 (prejudiced)	Total
Itagimirim	7	7	23	37
%	18,9%	18,9%	62,2%	100,0%
Itarantim	3	3	2	8
%	37,5%	37,5%	25,0%	100,0%
Itapebi	21	49	9	79
%	26,6%	62,0%	11,4%	100,0%
Salto da Divisa	10	27	84	121
%	8,3%	22,3%	69,4%	100,0%
Total	41	86	118	245
%	16,7%	35,1%	48,2%	100,0%

Source: Lima (2012)

Table 3. Distribution of groups according to the quality of public services questions, by *município*

	G1 (benefitted)	G2 (intermediary)	G3 (prejudiced)	Total
Itagimirim	16	20	15	51
%	31,4%	39,2%	29,4%	100,0%
Itarantim	7	1	7	15
%	46,7%	6,7%	46,7%	100,0%
Itapebi	31	26	16	73
%	42,5%	35,6%	21,9%	100,0%
Salto da Divisa	15	28	72	115
%	13,0%	24,3%	62,6%	100,0%
Total	69	75	110	254
Em %	27,2%	29,5%	43,3%	100,0%

Source: Lima (2012)

In terms of the socioeconomic groups, the ones that had their economic activities disturbed by the dam presented the most negative perception (fishermen, sand and stone extractors, laundresses), while landowners – who received compensation for the inundated land - were mostly in favor (Table 4).

Table 4. Distribution of groups according to socioeconomic groups

	G1 (benefitted)	G2 (intermediary)	G3 (prejudiced)	Total
Rock extractors	0	1	6	7
%	0,0%	14,3%	85,7%	100%
Sand extractors	0	0	4	4
%	0,0%	0,0%	100,0%	100%
Farmers	9	4	0	13
%	69,2%	30,8%	0,0%	100%
Washerwomen	0	1	4	5
%	0,0%	20,0%	80,0%	100%
Fishermen	0	3	11	14
%	0,0%	21,4%	78,6%	100%
Resettled households	0	10	7	17
%	0,0%	58,8%	41,2%	100%
Total	9	19	32	60
%	15,0%	31,7%	53,3%	100%

Source: Lima (2012)

Conclusion

This study shows that cluster analysis is a helpful to identify how differently social groups react to hydropower projects and, hence, can contribute to improve mitigation proposals during the license process and reduce social conflicts due to popular dissatisfaction with the projects.

In the case of the impacts of the Itapebi Hydroelectric Power Station, the cluster analysis showed that interviewees presented a significantly differentiation in their answers. Those who identified the most negative perception are the ones who live upstream, near the dam (especially in Salto da Divisa), or had their economic activities negatively affected by it (sand and rock extractors, washerwomen and fishermen). The people who felt mostly benefited are either located in the downstream *municípios*, or belong to the group of farmers and people who received new houses.

If the problems pointed out by those dissatisfied with the reservoir were anticipated before the construction, with a more robust and socially inclusive assessment of the impacts during the licensing period, a better project would have reduced, at least partially, the many conflicts that led to a large delay for the operation of the hydropower station.

The main conclusion is that, during the ex-ante impact assessment required for the licensing, cluster analysis based on interviews of the many diverse groups affected could be a very helpful tool to identify how differently the long term impacts are perceived, and more robust and less conflictive mitigation proposals can improve the social acceptance of the project.

References

CBEM, Centro Brasileiro de Energia e Mudanças Climáticas. **Desenvolvimento de metodologia de avaliação *ex post* dos impactos socioeconômicos e ambientais de empreendimentos hidrelétricos**. P&D ITA 05. Relatório Final. 2011.

GAVIAO, A. B. **Gestão de conflitos ambientais frente à implantação de hidrelétricas: estudo de caso do aproveitamento hidrelétrico de Itapebi/BA**. Dissertação de Mestrado em Regulação da Indústria de Energia. Universidade Salvador – UNIFACS. 2006.

KUBRUSLY, L. S. Um procedimento para calcular índices a partir de uma base de dados multivariados. **Revista de Pesquisa Operacional**, Vol. 21, Nº 1, p. 107-117, junho de 2001.

LIMA, G. R. 2012. **Análise dos impactos socioambientais de usinas hidrelétricas através do método de análise de grupamento.** Monografia de Graduação de Ciências Econômicas, Instituto de Economia, Universidade Federal do Rio de Janeiro, Rio de Janeiro.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2007.

YOUNG, C. E. F. ET al. **Metodologia de Avaliação E*x post* dos Impactos Econômicos e Socioambientais de Empreendimentos Hidrelétricos.** Anais do VI Congresso de Inovação Tecnológica em Energia Elétrica (VI CITENEL), Fortaleza/CE, 17 a 19 de agosto de 2011.