JOURNAL OF SYSTEMS
AND INDUSTRIAL PROJECT ENGINEERING
JSIPE

Vol.1 N°1, January 2015
pp. 10-17
ISSN :

Using Principal Component Analysis to Determine Key Factor of Rural Electrification Development Investment

Tokary A. Rajaoarisoa, François A. Ravalison
University of Antananarivo, Ecole Supérieure Polytechnique, 101 Antananarivo-Madagascar

ABSTRACT
The main objective of this study is to determine the key factors when decision on investment is required in rural electrification development. This study develops a process based on Principal Component Analysis to identify key factors. Principal Component Analysis is a data reduction method, that is, a method for reducing the number of variables. We have identified three variables related to investment in such context: State contribution, Bank contribution, Enterprise contribution and Commune contribution. We have extracted three factors: State contribution, Enterprise contribution and Commune contribution. In developing country and especially in the case of Madagascar, we have identified that State contribution plays a principal position in rural electrification development. The majority of rural electrification projects are Government contribution driven.

Keywords: *Investment, Principal Component Analysis, Rural Electrification Development*

1. INTRODUCTION

The majority of the Malagasy households live in inaccessible and scattered villages making the rural electrification through connection with the national network distribution complex. We know distributed energy systems, nowadays, are very hot subject all over the world [1]. Policy makers prefer decentralized energy solutions when connection with the national network is expensive. However, such preference requires hard technological solutions like the utilization of renewable energy. The primary advantage of the latter is its lack of greenhouse gas and other emissions in comparison with fossil fuel combustion [2]. The State policy is determinant and plays a great place as far as institutional and financial structure adoptions are suitable. Generally, the installation costs and production coupled with the faint need do not make profitable the offer of electric power. Related to this subject, we propose the following research question: what is the key factor when investment of a decentralized rural electrification project is concerned?

This study is structured as follows: a literature review on the general policy in the rural electrification program is presented. Then Principal Component Analysis is utilized to identify and visualize key factors. After this analysis, we can draw that State contribution in the investment holds a great part in the project. Then principal results will be listed. Furthermore, discussion will be conducted to answer the research question based on the results.

2. LITERATURE REVIEW

2.1 Rural electrification development

The decentralized rural electrification project is very expensive. Its operation cost is high but the rural people income is low. That makes the rural electrification difficult and the private investors discouraged. In this case, the project must be seriously subsidized to be viable [3] and sustainable. The emergence of energy sector reform such as structural changes, privatization, and relative success of the public-private partnership have given rise to solve those problems. Some countries stimulate the adoption of public-private partnership [4]. According to the analysis in industrialized countries, they mainly carried out reforms to facilitate the competition between suppliers, to improve the quality of service and to lower the prices of electricity. Moreover, in developing countries, they conducted reforms based on needs [5]. The target is to improve the technical and the financial execution on the first hand, and to support the investment in the system development and to reduce the policy interference in the tariff arrangement and the service management on the second hand [5]. In other terms, these reforms are based on market theories by which electricity is treated as product in opposition to old vision as an integrated service [6]. In these reforms, it is advisable to consider the difference between the entity offering the energy service and the owner of infrastructure. The ownership of infrastructure can has four possible cases:

- A public company ensured the electrical service. It is the most current form of the property, where the State must give an account of its management within the framework of contract-plan or contract-management.
- The State is owner, but the running operation is sub-contracted to the private sector.
- The private service is owner of the infrastructure but its activity is regulated.
- The Community ensures the service. This solution is privileged when the public service does not allow the satisfaction of such needs.

Our study is focused on the second case. The initial idea is to find a solution that encourages the investors to create a true appropriate business model related to profitable investments through security by institutions [7]. Then the State contribution concerns equipments. The private owner assures installation of these equipments. However, it is obvious that the public funding will continue to play an important part in rural electrification as far as the aim is to increase the accessibility rate of the rural people to electricity [8].

Besides, subscribers and technologies also play an important part in rural electrification. The financial assistance allocated for a project and the estimated numbers of subscribers are decisive factors the owners consider during invitation to tender. Choice of technology is important to determine the costs of investments and running operations [9]. At present, the evolution of renewable energy technology makes possible supporting the promotion of decentralized rural electrification policy [10]. More the environmental advantage brought by this technology is even more competitive, profitable and sustainable compared to other energy technology such as fossil energy [11]. Utilizable in these technologies are the photovoltaic solar system, the wind turbine, the hydroelectric and the gasification of biomass. Knowing that certain technology is expensive than the others for the same produced power, but it could be reconcilable with the processing site [12].

Indeed, rural electrification is complex without government intervening more or less energetically, because rural electrification is an expensive and unprofitable operation [13]. The developed countries knew a variety of institutional trajectories influenced by the institutional environment of each country. Moreover, electrification has found different translation according to countries. It clarifies the question of the complexity of electrification in the developing countries [14]. Among the biggest challenges is the regulatory design for achieving political independence and introducing rules to ensure accountability. There is a long way to follow in the best practices evolution for contracting, operating, and regulating rural electrification service concessions and in understanding the most appropriate standard of concessions in different contexts [15]. It seems whereas the speed with which electricity progressed in the villages the more the government intervenes, the more the rural people profit early from the electric service [16].

2.2 Principal Component Analysis (PCA)

The PCA is a method for multiple variables analysis, which seeks to identify the principal axes explaining the best correlation between descriptive variables [17] [18] [19]. These variables are characterized in a data matrix. The data processed by this method are presented in the statistical table. Let us note that a data table can be compared to a rectangular matrix. This matrix describes two different spaces as a vector space. A reading on line allows seeing the vectors where each element gives information on each variable for a given individual. A reading in column describes the variable vectors where each element gives the variation of the variable according to individuals of the sample.

In this table, the first column defines the individuals' identity. In order to obtain relevant results, it is necessary to analyze these individuals by eliminating the exceptional cases. The other columns of the table give the various variables that characterize the individuals. The following matrix $\boldsymbol{M_d}$ is a table of data with $\boldsymbol{N}$ observations and $\boldsymbol{k}$ variables:

$$M_d = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{Nk} \end{bmatrix} \quad (1)$$

$\boldsymbol{x_{ij}}$: indicate the variable i^{th} line and the j^{th}column.

For an initial data matrix with a dimension (***N;k***), the aims in this method are:

- To minimize the cloud points deformations, in the individuals space E_q with a dimension q $(q < N)$.
- To explain the best initial connections between these variables in the space F_q.
- To reduce the dimension of the initial matrix by a matrix of row q $(q < N)$.

The PCA interests the variables study in the table of information to confront the various distributions and to discover irregularities in these distributions. Besides, it proposes to analyze interrelationships between the variables, and to highlight the combinations between the variables.

The first stage of the method is the data matrix transformation into square matrix variance-covariance or correlation matrix. Since this method objective is to reduce the variables in order to determine the factorial axes or factors followed by a connection analysis or correlation between these variables. The statistical indicators that make this analysis possible are the variance covariance matrix $\mathbf{cov(X)}$ and the correlation matrix $\mathbf{cor(X)}$. For this matrix, the statistical concepts allow to write:

The variable average described in j^{th} column of the matrix M_d.

$$\bar{x}_{.j} = \frac{1}{N}\sum_{i=0}^{N} x_i \tag{2}$$

The variables covariance $\mathbf{x}_{.j}$ et $\mathbf{x}_{.l}$:

$$cov(x_{.j};x_{.l}) = \frac{1}{N}\sum_{i=0}^{N}(x_{ij}-\bar{x}_{.j})(x_{il}-\bar{x}_{.l}) \tag{3}$$

The variance of $\mathbf{x}_{.j}$:

$$var(x_{.j}) = \frac{1}{N}\sum_{i=0}^{N}(x_{ij}-\bar{x}_{.j})^2 \tag{4}$$

The standard deviation is the root of the variance:

$$\sigma_{.j} = \sqrt{\frac{1}{N}\sum_{i=0}^{N}(x_{ij}-\bar{x}_{.j})^2} \tag{5}$$

Let us transpose these formulas in the matrix case. We notice that the difference $x_{ij} - \bar{x}_{.j}$ explains the variation between the observation and the average value. The first transformation then consists in forming a square matrix $\mathbf{M_d}$ such as:

$$\mathrm{M_d} = \begin{bmatrix} x_{11}-\bar{x}_{.1} & x_{12}-\bar{x}_{.2} & \dots & x_{1k}-\bar{x}_{.k} \\ x_{21}-\bar{x}_{.1} & x_{22}-\bar{x}_{.2} & \dots & x_{2k}-\bar{x}_{.k} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ x_{N1}-\bar{x}_{.1} & x_{N2}-\bar{x}_{.2} & \dots & x_{Nk}-\bar{x}_{.k} \end{bmatrix} \tag{6}$$

As well as we have:

$$\mathrm{cov(X)} = \frac{1}{N}\mathrm{M_d^t}\ \mathrm{M_d} \tag{7}$$

$$\mathrm{cor(X)} = \left[\frac{cov(x_{.j};x_{.l})}{\sigma_{.j}\sigma_{.l}}\right] \tag{8}$$

Then, the second stage consists in determining the principal results of the method. Total inertia is one of them. Inertia measures the total cloud point dispersion. The principle is to obtain an under-space with a dimension on which projection resembles more with the initial cloud. In distance terms, in a vector space k dimension, the vector projection inertia in this under space p dimension lower than k is:

$$I(E_p) = \frac{1}{N}\sum_{i=1}^{N}(d(x_i;E_p))^2 \tag{9}$$

However, in algebraic term, the factorial plan is built starting from the resolution of an optimization problem. This plan results from the inertia projected cloud maximization. While passing by the theorem of the spectral decomposition, the under space E_p is the under vector space generated by the vectors $u_1;\dots;u_p$ associated with the eigenvalues $\alpha_1;\dots\alpha_p$ of the matrix $\mathbf{cov(X)}$ or the matrix $\mathbf{cor(X)}$.

The next result is the information-selected share. We carry out a change of reference in E_p in order to place a new representation system where the first axis as much as possible brings total inertia cloud. The second axis, as much as possible, centers inertia, which is not taken into account by the first axis, and so on. Therefore, as soon as we manage to determine the under space E_p we have all the factors. It may be that these factors are more than two. Then we consider only the plan generated by the eigenvectors associated by two greater Eigen values. When we consider more than two factors, we can couple these factors. The criterion to choose the factors is the following, which represents the information-selected share in the matrix k column after under space projection of dimension p lower than k if factors p are considered:

$$\frac{\sum_{i=1}^{p} \alpha_i}{\sum_{i=1}^{k} \alpha_i} \quad (10)$$

To obtain a plane representation of the individuals, it is necessary to build an orthogonal reference mark equipped with a standard value whose unit starts from the eigenvectors associated by two greater Eigen values. The cloud associated with the matrix is the orthogonal projection of each point. The coordinates of these projected points are called principal component noted c_i^j.

We can characterize those principal components. Let us consider the individual i located by his square cosine compared to the factorial axis. While noting θ the angle between the reference mark axis and the origin and the coordinate's individual distance, we have:

$$cos\theta = \frac{\sum_{i=1}^{p} (c_i^j)^2}{\sum_{i=1}^{k} (c_i^j)^2} \quad (11)$$

Besides, each individual has his contribution to the factorial axis formation. This indicator allows locating the aberrant value in the observations series. Indeed, the i individual contribution to the j axis formation is:

$$\frac{\frac{1}{N} c_i^j}{\alpha_j} \quad (12)$$

The orthogonal vectors define the factorial axes. In addition, orthogonal vectors are associated with the eigenvectors corresponding to the matrices $\mathbf{cov(X)}$ or $\mathbf{cor(X)}$ Eigen values. The variables projection on the factorial axis allows identifying the factorial axis significance. In other term, when we project the centered reduced variables on the factorial axis F_k, the projections belong to a circle of ray 1. The best representation is when the point is close to the circle. Therefore, the factorial charts that we have obtained previously concern only the individuals; that is to say variables. By using isometric transformation plan, we can simultaneously represent the individuals and the variables in only one factorial chart.

3. METHODOLOGY

To conduct the research, we have implemented the process below:

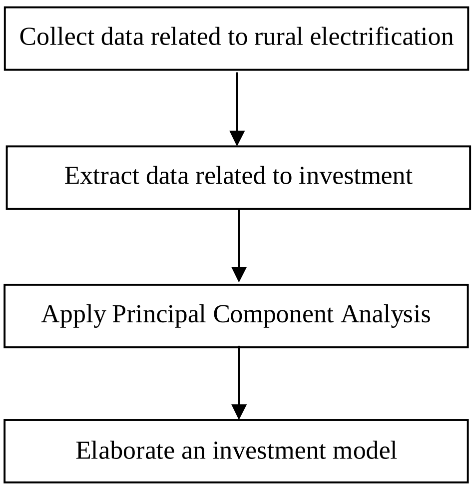

Figure 1: Research Process

We have collected data in 28 rural communes. The questions concern rural electrification. Then we have extracted data related to investment. We have applied Principal Component Analysis to those data.

4. RESULTS

When PCA is applied to the data table with Pearson Correlation Coefficient (variance with 1/n). We have obtained the correlation matrix in table 1.

Table 1: Correlation matrix

	Government	Enterprise	Community	Bank
Government	1	**0.716**	-0.082	-0.114
Enterprise	**0.716**	1	-0.100	-0.072
Community	-0.082	-0.100	1	-0.057
Bank	-0.114	-0.072	-0.057	1

In bold, significant values (0.716) at the level of significance alpha=0.050 (two-tailed test) are obtained. Government and enterprise are correlated.

Then we consider the plan generated by the eigenvectors associated by two greater Eigen values. The table 2 below shows those Eigen vectors.

Table 2: Eigen vectors

	F1	F2	F3	F4
Government	0.691	0.025	0.142	0.708
Enterprise	0.688	-0.029	0.170	-0.705
Community	-0.154	0.707	0.690	-0.013
Bank	-0.158	-0.706	0.689	0.041

In axis F1, Government and enterprise Eigen vectors are respectively 0.691 and 0.688. We notice that Government Eigen value is slightly superior to enterprise one.

By using isometric transformation plan, the following figure 2 represents the individuals and the variables in only one factorial chart.

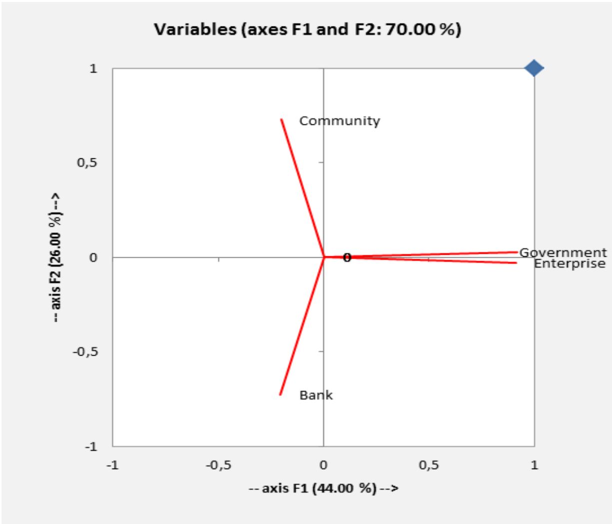

Figure 2: Principal component analysis

In axis F1, we observe that Government contribution and enterprise contribution are correlated.

The following table 3 shows the contributions of the variables.

Table 3: Contributions of the variables (%)

	F1	F2	F3	F4
Government	47.759	0.062	2.012	50.167
Enterprise	47.379	0.086	2.888	49.647
Community	2.364	49.992	47.627	0.017
Bank	2.498	49.860	47.473	0.169

At F1 axis level, contributions of Government and Enterprise are respectively 47.75% and 47.37%.

5. DISCUSSION

We have four variables or four types of financial contributions in this study: Government, Enterprise, Community and Bank. In the correlation matrix, the linear relation between the Government and Enterprise variables are significant. The negative sign of the linear correlation coefficient between these variables results in the variation, which is done in the opposite direction. In other words, if the Government financial contribution increases, this involves a reduction in the Enterprise financial contribution and conversely. In the same way when the Commune financial contribution decreases then the Government financial contribution increases.

Concerning the variables, we notice the three following financial contributions: Government, Enterprise and Community that represent the best in the factorial chart formed by the F1 and F2 axes because they are in the quadrants [0,1]. Then, the position of each variable justifies the complementarities between these values: the variables Community and Government are in quadrant [0,1] of F2 axis. In addition, variables Government and Enterprise are in quadrant [0,1] of F1 axis. Financial contributions of Government influence financial contributions of Community and Enterprise. We also view this connection between the Community and Enterprise variables with the Government variable such as the coefficient correlation preceded by a negative sign. Lastly, the variables square cosine compared to the factorial chart axes shows that the Government variable dominates the F1 axis while the Community variable orders the F2 axis. Thus, the F2 positive part shows the strong Community contribution and the negative part proves the importance of Enterprise contribution. In short, the Government share in a rural electrification project investment is dominant.

According to the literature review, Government should bring more to carry out a rural electrification project. The results above confirm that Government part occupies an important place in the corresponding investment. Despite the initiative takes by Government, access to rural electrification is still in a low level. That is why the vision «energy for all» has become all developing countries' vision. The cost of the project is the principal problem. For the case that we are dealing here, private enterprise holds the operation contract with the Government in order to manage the electrical energy supply of a Community. Of course, after analyzing the business plan of those enterprises we can deduce that their part on investment associated with their operation cost influence the bills the subscribers have to pay. Thus, the rural electrification achievement is shaped by a triangle enterprise profit/populations' ability to pay/capacity of grantor authority to subsidize investment. However, given that enterprise is looking for an acceptable profit, considering also that electrical service costs highly exceed population abilities to pay, grantor authority generally agree to allocate financial support to investment so that it can merge users and enterprise's interests. A new form of partnership needs innovation process so that project will be viable for all stakeholders.

Nevertheless, prediction of electricity consumption plays an important role in future energy related research and planning strategies [20]. What is more, in an adequate model of partnership, Enterprise will always stay directing the work and running the equipment during the contract period. This idea is far being satisfactory to stop the profit of private operators but interesting in the competitiveness sector. However, a unit should be set up to analyze activities and local opportunities for the eventual program using at the first time an electrical energy. It will be in charge of the project launching and of different missions, as well as the selling of the energy and the promotion of the program linked to a social, economical, or environmental project. That is why it will have the idea and the instruments of production modernization in rural field. Then it will play the interface between users and Government. In this case, it can be the new partner in the investment process. This idea will reduce fixed cost effect on price. Because price is an important factor although capacity utilization is more important in electricity prediction [20]. Moreover, technological choice is an influential factor in investment.

6. CONCLUSION

Rural electrification project faces many problems because fossil resources are distinctly expensive [21]. Besides, distributed energy systems are very hot subject all over the world [1] and above all in Madagascar. Government has to intervene so that rural population accessibility to electricity is feasible and effective. Such intervention should be coupled with enterprise's intervention as far as running operation is concerned.

Bank contribution is relatively impossible. Bank process is not suitable for rural context, rate of return is low and investment return takes a long time. So rarely, rural electrification projects are bank financed. In the factorial chart, Bank contribution is in negative quadrant for both F1 and F2.

Considering the above problems, the movement from small, pilot, and experimental renewable energy projects is a must and inevitable [2]. Alsaqoor and al. confirm such concept when they say the challenge is on extensive use of renewable sources [1].

There are some domains for further research when rural electrification development is concerned:

- How could Community become a key factor in electrification development?
- What process are we going to implement to popularize the use of renewable sources?

7. REFERENCES

[1] S. Alsaqoor, M. AlAjlouni, K. AlQdah, H. Kruczek, E. Pelinska-Olko, "Development Chances of Distributed Energy Production on Small Scale", Jordan Journal of Mechanical and Industrial Engineering, Vol.4, N°1, 2010, 135-142

[2] M. Al zou'bi, "Renewable Energy Potential and Characteristics in Jordan", Jordan Journal of Mechanical and Industrial Engineering, Vol.4, N°1, 2010, 45-48

[3] Davis M., "Institutional Frameworks for Electricity Supply to Rural Communities", EDRC, 1995

[4] N. Wamukonya, "Power Sector Reforms in Developing Countries: Mismatched Agendas", in Wamukonya Njeri, eds. Electricity Reforms: Social and Environmental Challenges, UNEP Collaborating Centre on Energy and Development, Roskilde, Denmark, 2003

[5] Bhagavan M.R. Reforming the Power Sector in Africa, African Energy Policy Research Network (AFREPREN), Zed Books Limited, London, United Kingdom & New York, USA, 1999

[6] J. Byrne, Y. Mun, "Rethinking Reform in the Electricity Sector: Power Liberalization or Energy Transformation", in Njeri Wamukonya, eds. Electric Reform: Social and Environmental Challenges, UNEP Riso Centre, Roskilde, Denmark, 2003

[7] A. Jadresic, "Promoting Investment in Rural Electrification, the Case of Chile", the World Bank Viewpoint, N° 214, 2000

[8] M. Nangammbi, L. Mosomane, "Rural Electrification In The Energy Policies Of The SADC Region", Sustainable Energy Futures Group-Natural Resources And The Environment, 2011

[9] E. Martinot, R. Kilian, "Regulatory Approaches to Rural Electrification and Renewable Energy: Case Studies from Six Developing Countries", the World Bank-Washington DC, Working Paper, June 2000

[10] Haanyika M.C.; "Rural Electrification Policy And Institutions In A Reforming Power Sector"; In Energy Policy Journal, Elsevier Ltd; October 2005

[11] S. Kaufman, "Rural Electrification with Solar Energy as a Climate Protection Strategy", Renewable Energy Policy Project Research Report No. 9, 2000

[12] ESMAP, Energy and Development Report 2000, Energy for the World's Poor, the World Bank, Washington DC, 2000

[13] Brown D.C. Electricity for Rural America: the Fight for REA, Westport, Greenwood Press, 1980

[14] Finon D. La Variété des Trajectoires Institutionnelles d'Electrification, Working paper CIRED (Centre international de recherche sur l'environnement et le développement), 2006

[15] A. Estache, "Designing Regulatory Institutions for Infrastructure Lessons from Argentina", the World Bank Public Policy Note N°114, 1997

[16] Y. Tremblay, "Les Etapes de l'Electrification Rurale dans l'Ouest Canadien, 1920-1960", Cahiers franco-canadiens de l'ouest, Vol.7, N°1, 1995, 37-67

[17] Jaeon K., Mueller C.W, "Introduction to Factor analysis: What it is and how to do it", Beverly Hills, Sage Publications, 1978, 80-160

[18] J. Shlens, "A Tutorial on Principal Component Analysis", Systems Neurobiology Laboratory, Salk Institute for Biological Studies and Institute for Nonlinear Science, University of California, San Diego, Version 2, 2005
[19] Steven M. H. Principal Components Analysis, University of Georgia, Athens, 2008
[20] A. Al-Ghandoor, M. Samhouri, "Electricity Consumption in the Industrial Sector of Jordan: Application of Multivariate Linear Regression and Adaptive Neuro-Fuzzy Techniques", Jordan Journal of Mechanical and Industrial Engineering, Vol.3, N°1, 2009, 69-76
[21] O. Badran, E. Abdulhadi, R. Mamlook, "Evaluation of Solar Electric Power", Jordan Journal of Mechanical and Industrial Engineering, Vol.4, N°1, 2010, 121-128