Principal component regression analysis of electricity consumption factors in Madagascar

Ramaharo, Franck M. and Rajaonarison, Njakanasandratra R.

26 January 2023

Principal component regression analysis of electricity consumption factors in Madagascar

Franck M. Ramaharo[1] and Njakanasandratra R. Rajaonarison[2]

[1]Service de la Modélisation Économique
Ministère de l'Économie et des Finances
Antananarivo 101, Madagascar
franck.ramaharo@gmail.com

[1]Mention Économie
Université d'Antananarivo
Antananarivo 101, Madagascar
njakanasandratra.rajaonarison@gmail.com

January 26, 2023

Abstract

We carry out principal component regression analysis on a dataset consisting of selected relevant economic and energy indicators in order to identify the main driving factors impacting the electricity consumption in Madagascar. Our results show that in accordance with the country's current energy profile, factors related with rural electrification have the most influence on electricity consumption in Madagascar.

Keywords: principal component analysis, principal component regression, electricity consumption.

1 Introduction

Electricity consumption is considered as a key indicator of economic development. When calculated per capita, electricity consumption in advanced economies is significantly higher than those of developing economies. As for 2019, the electricity consumption in Madagascar is only 74.1 kWh per capita, which is about 21.5% of the Sub-Saharan

Africa average, and compared, for instance, to 11,894.5 kWh per capita in Northern America [17, p. 56, 61]. In fact, this is one of the lowest electricity consumption in the world. To better understand this issue, we need to identify the driving factors impacting the electricity consumption in Madagascar. Such identification is conducted by applying principal component analysis to a dataset consisting of selected relevant economic and energy indicators in the country (§§3.1). Then, we fit a multiple linear regression model on the retained principal components in order to assess relationships between electricity consumption and its driving factors (§§3.2) and to determine the influence of each factor on electricity consumption (§4).

2 Madagascar energy profile

The most dominant energy source in Madagascar consists of solid biomass (see Figure 1). As for 2017, electricity only contributes to 0.9% of the total energy supply.

Around 43.2% and 2.0% of the electricity production in 2021 delivered by the JIRAMA (Jiro sy Rano Malagasy, the national electricity company of Madagascar) are supplied by hydroelectric stations and other renewable energy sources (RES), which include solar photovoltaic, wind energy and biomass, respectively [1, p. 53]. The remaining 54.8% is secured by thermal power stations running on imported heavy fuel oil and diesel fuel (see Figure 1). It is worth mentioning that energy imports, which consist mostly of petroleum products, account for 13.1% of the total imports of goods in 2021 [2, p. 108].

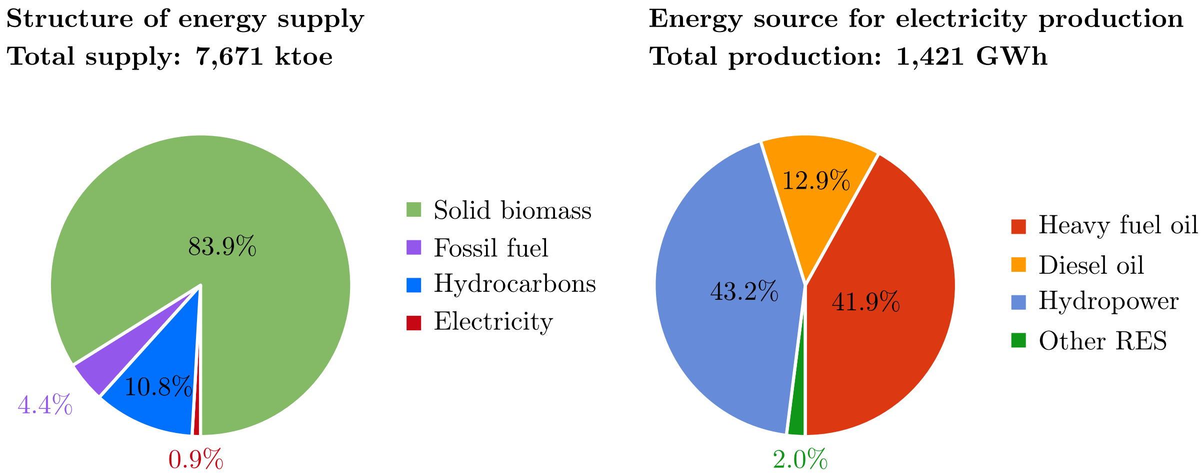

Figure 1: Madagascar energy profile [1, 10].

Electricity consumers in Madagascar are classified into four categories: residential, industrial, small and medium enterprises/services, and public lighting [13]. Total electricity consumption reached 1,372.7 GWh in 2021 (compared to 1,333.6 GWh in 2020), of which 51.3%, 48.3% and 0.4% are attributed to households, industries and services, and public lighting, respectively [8, p. 49].

The electrification rate in Madagascar is among the lowest in Africa. In 2021, only 16.6% of the population have access to modern electricity (compared to 16.4% in 2019). Rural areas of the country, where lower-income households tend to live, are unequally electrified with an approximate electrification rate of just 6.5% [7]. Households in rural areas (which represent 78.9% of all households [4, p. 39]) use fuelwood and charcoal for heating and cooking. Due to lack of affordability and accessibility, lower-income households rely on kerosene lamps, candles, batteries and diesel generators which are mostly limited to lighting [12]. The Ministry of Energy and Hydrocarbons estimates that households with no access to electricity spend an average of 6.1% of their income on energy [11, p. 52].

Installed capacity for electricity production in the country is low and the distribution networks are poorly developed to meet the growing demand. In addition, although electricity prices are too high (one of the highest among the SADC region members [14], and one of the least affordable to households in Sub-Saharan Africa [6]), they do not cover the production costs. Moreover, recurring power cuts and rising energy prices (note that the share of energy item in the new consumer price index is 0.2 [3]) have repercussion on economic activities (industrial sector and service sector contribute approximately 15.0% and 57.2% to overall gross domestic product, respectively [5]).

3 Methodology

Based on literature [13, 15, 23, 24] and data availability, we select 9 predictor variables: industry net output ($\boldsymbol{x}_1$), service net output ($\boldsymbol{x}_2$), gross domestic product – GDP ($\boldsymbol{x}_3$), gross disposable private income ($\boldsymbol{x}_4$), consumer price index ($\boldsymbol{x}_5$), percentage of rural population ($\boldsymbol{x}_6$), energy imports ($\boldsymbol{x}_7$), electricity generation ($\boldsymbol{x}_8$), access to electricity ($\boldsymbol{x}_9$). The dependent variable is electricity consumption ($\boldsymbol{y}$). Data are retrieved from the World Bank Database [20, 21], the National Institute of Statistics [5] and the U.S. Energy Information Administration [18, 19]. For each variable, we have data from 1992–2020, and for the purpose of the analysis, all variables are already standardized. Principal component analysis is then applied to the dataset.

3.1 Principal component analysis

Principal component analysis is a statistical tool for reducing the dimensionality of a dataset consisting of a large number of interrelated variables into new uncorrelated variables, while retaining the maximum of the variability present in the dataset [16]. The new uncorrelated variables are referred to as *principal components* and each of them is a linear combination of the original variables.

Given the original dataset $(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_N)$, we first compute the covariance matrix in order to reveal the linear correlations between the variables (see Table 1). The covariance matrix exhibits a number of high positive (0.9997) and negative (-0.9987) correlations. Notice that $\boldsymbol{x}_6$ is highly correlated in a negative manner with all the other variables, meaning that the percentage of population, the economic indicators $(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4, \boldsymbol{x}_5)$ and the energy distribution indicators $(\boldsymbol{x}_7, \boldsymbol{x}_8, \boldsymbol{x}_9)$ tend to move in opposite directions. This expresses how the great majority of Malagasy rural population is mainly dependent on subsistence farming and natural resource extraction which remain the principal economic activities providing employment, household incomes and energy needs.

The next step is to perform the eigendecomposition of the covariance matrix. Recall that the eigenvectors of the covariance matrix represent the directions of the axes (the

	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$	$\boldsymbol{x}_4$	$\boldsymbol{x}_5$	$\boldsymbol{x}_6$	$\boldsymbol{x}_7$	$\boldsymbol{x}_8$	$\boldsymbol{x}_9$
$\boldsymbol{x}_1$	1.0000								
$\boldsymbol{x}_2$	0.9902	1.0000							
$\boldsymbol{x}_3$	0.9908	**0.9997**	1.0000						
$\boldsymbol{x}_4$	0.9873	0.9972	0.9971	1.0000					
$\boldsymbol{x}_5$	0.9720	0.9929	0.9936	0.9915	1.0000				
$\boldsymbol{x}_6$	-0.9659	-0.9894	-0.9899	-0.9886	**-0.9987**	1.0000			
$\boldsymbol{x}_7$	0.7646	0.8027	0.8032	0.8014	0.8387	-0.8486	1.0000		
$\boldsymbol{x}_8$	0.9410	0.9593	0.9601	0.9590	0.9710	-0.9771	0.8603	1.0000	
$\boldsymbol{x}_9$	0.9035	0.8867	0.8915	0.8848	0.8694	-0.8571	0.6077	0.8317	1.0000

Table 1: Correlation coefficients of predictor variables.

principal components), and where there is the most variance. Besides, eigenvalue represents the total amount of variance that can be explained by a given principal component. In Table 2, we can see that the first component accounts for 93.0326% of the total variability in the original variables, while the second and third component only accounts for 4.753% and 1.334%, respectively. We retain the first two principal components which carry a cumulative 97.780% of the total variability.

Component	Initial eigenvalues			Extraction sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	8.372	93.026	93.026	8.372	93.026	93.026
2	0.428	4.753	97.780	0.428	4.753	97.780
3	0.120	1.334	99.114			
4	0.048	0.536	99.650			
5	0.025	0.274	99.924			
6	0.004	0.049	99.974			
7	0.002	0.019	99.993			
8	0.001	0.007	99.999			
9	0.000	0.001	100.000			

Extraction method: Principal Component Analysis.

Table 2: Total variance explained.

These two principal components are created as linear combinations of the original variables. The coefficients that are used to create the linear combinations are given by the *feature vector* (Table 3). The first principal component has a nearly equal positive load on variables $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, $\boldsymbol{x}_4$, $\boldsymbol{x}_5$, $\boldsymbol{x}_8$, $\boldsymbol{x}_9$ and of opposite sign on $\boldsymbol{x}_6$, while the second principal component has a larger positive load on $\boldsymbol{x}_7$. We may interpret the first principal component as representative of electricity availability and economic activity in Madagascar, and the second one as representative of energy import dependency.

	$\boldsymbol{e}_1$	$\boldsymbol{e}_2$
Industry net output ($\boldsymbol{x}_1$)	**0.3398**	-0.1784
Service net output ($\boldsymbol{x}_2$)	**0.3439**	-0.0762
Gross domestic product ($\boldsymbol{x}_3$)	**0.3442**	-0.0813
Gross disposable private income ($\boldsymbol{x}_4$)	**0.3434**	-0.0747
Consumer price index ($\boldsymbol{x}_5$)	**0.3441**	0.0292
Percentage of rural population ($\boldsymbol{x}_6$)	**-0.3436**	-0.0701
Energy imports ($\boldsymbol{x}_7$)	0.2910	**0.7797**
Electricity generation ($\boldsymbol{x}_8$)	**0.3372**	0.1539
Access to electricity ($\boldsymbol{x}_9$)	**0.3084**	-0.5592

Table 3: Feature vector.

3.2 Principal component regression

Principal component regression (PCR) is a regression analysis technique that is based on principal component analysis. In PCR, we fit a linear regression model on the k selected principal components instead of regressing the dependent variable on the predictor variables directly.

Let $\boldsymbol{z}_i$ denote the i-th principal component, and let k denote the number of selected principal components. We have

$$\boldsymbol{z}_i = \sum_{j=1}^{N} \boldsymbol{e}_{i,j} \boldsymbol{x}_j, \quad i = 1, \ldots, k. \tag{1}$$

Then, the PCR model is

$$\widehat{\boldsymbol{y}} = \sum_{i=1}^{k} \alpha_i \boldsymbol{z}_i, \tag{2}$$

where $\widehat{\boldsymbol{y}}$ is an estimate of the dependent variable $\boldsymbol{y}$. By (1), the final model is obtained by transforming the parameter coefficient estimates back to the scale of original variables:

$$\widehat{\boldsymbol{y}} = \sum_{i=1}^{k} \alpha_i \left(\sum_{j=1}^{N} \boldsymbol{e}_{i,j} \boldsymbol{x}_j \right) = \sum_{j=1}^{N} \left(\sum_{i=1}^{k} \alpha_i \boldsymbol{e}_{i,j} \right) \boldsymbol{x}_j. \tag{3}$$

4 Results and discussion

Table 4 show the result of the PCR. The fitted regression model is

$$\widehat{\boldsymbol{y}} = 0.336 \boldsymbol{z}_1 + 0.162 \boldsymbol{z}_2. \tag{4}$$

The overall regression is statistically significant ($\mathrm{R}^2 = 0.959$, $\mathrm{F}(2, 27) = 314.818$, $p < 0.01$).

By (3), the final model is

$$\begin{aligned} \widehat{\boldsymbol{y}} = 0.085\boldsymbol{x}_1 + 0.103\boldsymbol{x}_2 + 0.103\boldsymbol{x}_3 + 0.103\boldsymbol{x}_4 + 0.121\boldsymbol{x}_5 \\ -0.127\boldsymbol{x}_6 + 0.225\boldsymbol{x}_7 + 0.138\boldsymbol{x}_8 + 0.013\boldsymbol{x}_9. \end{aligned} \tag{5}$$

	Dependent variable:
	$\boldsymbol{y}$
$\boldsymbol{z}_1$	0.336^{***}
	(0.013)
$\boldsymbol{z}_2$	0.162^{**}
	(0.060)
Observations	29
R^2	0.959
Adjusted R^2	0.956
Residual Std. Error	0.206 (df = 27)
F Statistic	314.818^{***} (df = 2; 27)
Note:	$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table 4: Regression result.

It can be seen that except for the percentage of rural population ($\boldsymbol{x}_6$), all variables have positive impact on electricity consumption. An increase, by one unit, of $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, $\boldsymbol{x}_4$, $\boldsymbol{x}_5$, $\boldsymbol{x}_7$, $\boldsymbol{x}_8$ and $\boldsymbol{x}_9$ would result in a rise in electricity consumption by 0.085, 0.103, 0.103, 0.103, 0.121, 0.225, 0.138 and 0.013 units, respectively. On the other hand, a rise of $\boldsymbol{x}_6$ by one unit brings a decline in electricity consumption by 0.127.

Regarding the positive coefficients in (5), we have the following remarks.

(i) The sectors of industry and services were responsible for around 48.3% of the final electricity consumed in 2021 in Madagascar. Therefore, it is clear that an increase in industrial and service activities would stimulate electricity demand. Besides, the model also tells us the country's electricity consumption tends to increase along with the development of productive economic activities, here represented by its GDP.

(ii) We might expect that electricity consumption decreases when prices increase. But what the model tells us is that electricity consumption can increase while prices increase, provided that income also increases.

(iii) The coefficient attached to $\boldsymbol{x}_7$ (energy imports) is the largest in our model (0.225). Since Madagascar's current electricity generation relies heavily on imported oil and diesel sources, this coefficient reflects the extent to which the economy of Madagascar depends on imports in order to meet its energy needs.

(iv) Madagascar's electricity generation can be improved, not only by developing and expanding the network of small hydroelectric power plants, but also by exploiting the large existing potential of renewable energy sources (Madagascar is already aiming for 85% of power generation to come from renewables by 2030 [9, p. 10]).

(v) Increasing access to electricity can be achieved through grid extension and interconnections, as well as off-grid renewable energy solutions, such as solar home systems and mini-grids.

Finally, the negative coefficient of $\boldsymbol{x}_6$ tells us that an increase of the percentage of rural population implies a decrease of electricity demand since a large majority of rural population relies heavily on biomass and do not have access to modern energy. Hence, electrification should be accelerated through grid extension in rural and remote areas. Access to electricity should be facilitated by promoting the supply of electricity services so that rural households could have access to a minimum level of electricity, and by promoting the widespread adoption and sustainable use of all forms of renewable energy.

References

[1] Projet ARCEB (2022). *Analyse approfondie de l'ensemble des contraintes qui freinent le développement du secteur de l'électricité*, Ministère de l'Énergie et des Hydrocarbures - Projet ARCEB, Madagascar.

[2] Banky Foiben'i Madagascar (2021). *Rapport Annuel 2021*, Banque Centrale de Madagascar: Direction de la Documentation et des Centrales d'Information, Madagascar.

[3] Institut National de la Statistique (2020). Le nouvel indice des prix à la consommation, INSTAT: Direction des Statistiques des Conditions de Vie des Ménages, Madagascar.

[4] Institut National de la Statistique (2020). *Résultats Globaux du Recensement Général de la Population de l'Habitation de 2018 de Madagascar (RGPH-3)*, Tome 1, Ministère de l'Économie et des Finances, Madagascar.

[5] Institut National de la Statistique. Comptes nationaux, `https://www.instat.mg/statistiques/bases-de-donnees/comptes-nationaux` [accessed 27 April 2022].

[6] Masami Kojima, Xin Zhou, Jace Jeesun Han, Joeri de Wit, Robert Bacon and Chris Trimble (2016). Who uses electricity in sub-Saharan Africa? Findings from household surveys, *Policy Research Working Paper*, No. **7789**, Washington D.C.: World Bank Group.

[7] Jon Lan, Micol Salmeri, Benjamin Curnier and Yuri L. Hamden (2019). *Mini-Grid Market Opportunity Assessment: Madagascar*, Green Mini-Grids Market Development Programme Document Series, African Development Bank Group.

[8] Ministère de l'Économie et des Finances (2022). *Rapport de mise en œuvre des programmes du gouvernement, 2021*, Ministère de l'Économie et des Finances: Direction de la Coordination et du Suivi Évaluation des Programmes, Madagascar.

[9] Ministère de l'Énergie et des Hydrocarbures (2015). *Lettre de Politique de l'Énergie de Madagascar 2015–2030*, Ministère de l'Énergie et des Hydrocarbures, Madagascar.

[10] Ministère de l'Énergie et des Hydrocarbures (2019). *Bilan Énergique National 2017*, Ministère de l'Énergie et des Hydrocarbures, Madagascar.

[11] Ministère de l'Énergie et des Hydrocarbures (2018). *Analyse des options d'électrification géospatiale au moindre coût pour un déploiement sur réseau et hors réseau*, Ministère de l'Énergie et des Hydrocarbures, Madagascar.

[12] Kameshnee Naidoo and Christiaan Loot (2020). *Madagascar. Energy and the Poor: Unpacking the Investment Case for Clean Energy*, UN Capital Development Fund (UNCDF).

[13] Jean Philippe Praene, Rindrasoa M. Rasamoelina and Leslie Ayagapin (2021). Past and prospective electricity scenarios in Madagascar: The role of government energy policies, *Renewable and Sustainable Energy Reviews*, Volume **149**, 111321.

[14] Ketakandriana Rafitoson (2017). *La lente marche vers la transition énergétique à Madagascar: État des lieux et perspectives*, Friedrich-Ebert-Stiftung, Madagascar.

[15] Tokary A. Rajaoarisoa and François A. Ravalison (2015). Using principal component analysis to determine key factor of rural electrification development investment, *Journal of Systems and Industrial Project Engineering*, Volume **1**(1), 10–17.

[16] Lindsay I. Smith (2002). *A Tutorial on Principal Components Analysis*, University of Otago, New Zealand.

[17] United Nations Statistics Division (2022). *2022 Energy Statistics Pocketbook*, United Nations, New York.

[18] U.S. Energy Information Administration. Electricity consumption, `https://www.eia.gov/international/data/country/MDG/electricity/electricity-consumption` [accessed 01 Dec. 2022].

[19] U.S. Energy Information Administration. Electricity generation, `https://www.eia.gov/international/data/country/MDG/electricity/electricity-generation` [accessed 01 Dec. 2022].

[20] World Bank. Access to electricity (% of population) - Madagascar, `https://data.worldbank.org/indicator/EG.ELC.ACCS.ZS?locations=MG` [accessed 01 Dec. 2022].

[21] World Bank. Rural population (% of total population) - Madagascar, `https://data.worldbank.org/indicator/SP.RUR.TOTL.ZS?locations=MG` [accessed 01 Dec. 2022].

[22] World Bank (2017). *Small hydro resource mapping in Madagascar: Executive summary : phase 3 - production of a validated resource atlas*. Energy Sector Management Assistance Program Washington, D.C.: World Bank Group.

[23] Li Yingying and Niu Dongxiao (2010). Application of principal component regression analysis in power load forecasting for medium and long term, *2010 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*, pp. V3-201-V3-203.

[24] Jing Zhang, Xin-yao Yang, Fei Shen, Yuan-wei Li, Hong Xiao, Hui Qi, Hong Peng and Shi-huai Deng (2012). Principal component analysis of electricity consumption factors in China, *Energy Procedia*, Volume **16**, 1913–1918.