International Journal of
Sustainable Energy Planning and Management

Modelling of rural electrical appliances ownership in developing countries to project their electricity demand: A case study of sub-Saharan Africa

Cristina Dominguez[a,b]*, Kristina Orehounig[b] and Jan Carmeliet[a]

[a] *Chair of Building Physics, Swiss Federal Institute of Technology Zürich (ETHZ), Stefano-Franscini-Platz 1, 8093 Zürich, Switzerland*
[b] *Laboratory for Urban Energy Systems, Swiss Federal Laboratories for Materials Science and Technology (Empa), Überlandstrasse 129, 8600 Dübendorf, Switzerland*

ABSTRACT:

Rural electrification projects are mostly listed as a priority in developing countries due to the numerous benefits it brings to improve people's quality of life. The starting point for these projects is the accurate estimation of the households' energy demand, but the large amount of resources that requires the on-site data collection makes this process unattractive to investors. This paper brings a potential solution to that matter, presenting a methodology for modelling the appliances ownership of rural households in developing countries to project their electricity demand. Based on a statistical approach, and including training data from more than 1,100 household samples from Nigeria and Ethiopia, the correlation between household survey data of ownership of the most common electricity-consuming appliances in developing countries and different socio-economic, demographic and geographic variables are investigated. Its accuracy is tested using other sets of validation data at a household, state, and national level. Finally, the results are projected using a Geographic Information Systems approach for identifying possible sites to be electrified. The high potential of using this approach for projecting the appliances ownership for rural areas was proven by obtaining a good overall accuracy in the results. As expected, the errors obtained at a household level were bigger than at a state or national level, these deviations are attributed to the presence of outliers due to the human behaviour incidence at household scale; which also affects the correlation patterns between the appliances ownership and evident socioeconomic factors.

Keywords:

Appliances ownership;
Energy demand modelling;
Geographic Information Systems;
Rural electrification;

URL:
http://dx.doi.org/10.5278/ijsepm.2564

1. Introduction

Around 17% of the global population lacks access to electricity, 84% of this population is located in rural areas of developing countries; from which, more than 95% is in sub-Saharan Africa [1]. The relationship between rural electrification and development has been recognized by different studies, emphasising on the improvement of the economic productivity, among other health and education benefits [2-4]. Consequently, rural electrification projects are mostly considered as a top priority for the national authorities in developing countries. To improve the planning of these projects, it is essential to have an accurate knowledge of the current electricity consumption of those areas, but not having records and explicit databases makes this a challenging task.

While rural areas are characterised by having low income per capita and low energy consumption [5], national household surveys present evidence of their electrical appliances ownership even when they have limited or no electricity access at all. This fact gives a

*Corresponding author - e-mail: dominguc@ethz.ch

sight of the potential of using these data to estimate their electricity demand. Different studies have been done regarding the estimation of energy consumption in developing countries. Some have included an analysis of fuels consumption patterns in developing countries, as the authors in [6] did with a time series analysis for Ghana, Kenya and South Africa; other studies have focused on studying the energy consumption in residential buildings based on heating and cooling energy needs accounting for climatic constraints, as the authors in [7] for Algeria. Other authors, such as [8], have presented methods to model the load profiles of rural mini and micro-grids considering as one of the inputs the total number of appliances. Conversely, scarce amounts of researchers have studied bottom-up methodologies to model the energy consumption based on the diffusion of appliances in the residential sector at a national level. For example, in [8] and [9] the authors consider macroeconomic variables such as income per capita, national urbanization and electrification rates in order to give an estimation of the diffusion of some electrical appliances in developing countries over time applying logistic regression. However, these macroeconomic drivers do not represent the households' conditions in rural areas, ensuing an overestimation of the results when it comes to modelling their electricity demand. Other studies have focused on modelling the trend of acquisition of electrical appliances based on historical data of household-level determinants in developing countries, such as in [10] for Brazil, India and South Africa. In the mentioned study, the authors use a machine learning algorithm (Boosted Regression Trees) and logistic regression in order to analyse the acquisition trends in these countries at a national level dividing the population on income quantiles. Nonetheless, they only consider high-power-consuming appliances (refrigerators, televisions, and washing machines) which most of the time only the population connected to the national grid can afford. In addition, none of these studies has included a geographical analysis to identify the differences of appliances ownership within the countries, and they are only limited to make comparisons with other countries at a national scale.

The methodology proposed in this paper aims at filling in this research gap by applying a statistical and high spatial resolution data processing approach. The first step of the analysis is focused on gathering the most recent available data of appliances ownership and different indicators of living standards of rural areas in developing countries; these were obtained by the Living Standard Measurement Study (LSMS) held by the World Bank [11]. It contains information at a household level that aims to support policymakers in developing countries; while it is not focused on the energy sector, it contains detailed data from a broad geographic scope that can be used for this research purpose. The collected data is then correlated using the multiple linear regression (MLR) method, creating a model for each of the four appliances chosen for this study (televisions, radios, mobile phones, and refrigerators). These models depend on different variables, such as household size, the proximity of the household to a major road, and electricity access; as well as specific information of the household head (age, literacy, religion, and employment situation). An analysis of the models' selection and the variables used is presented, followed by the measurement of the models' performance at different spatial scale. The results are projected using Geographic Information Systems (GIS) country maps and projections of the existing and planned electricity grid as an example of finding the potential areas to be listed as priority for future electrification projects.

1.1. Case study areas

Being sub-Saharan Africa the region with the lowest electrification rates of the world [1], it was of particular interest to focus this research work on it. The selected countries from the region were Nigeria and Ethiopia due to the availability of recent data from the LSMS household surveys [11], which cover the years 2015 – 2016. These countries are also considered as an interesting reference because of their differences in macroeconomic aspects and their similarities in natural resources (Table 1). It is well known that rural areas depend mostly on agriculture and related activities for their livelihoods [11]; also, that appliances acquisition is related to their income [10]. The International Institute for Applied Systems Analysis (IIASA) and the Food and

Table 1. Demographic, economic and geographic characteristics, data for 2016 [1,14,15]

Indicator	Nigeria	Ethiopia
Population, millions	185.99	102.40
GDP, billion US$	404.65	72.37
Rural electrification rate, %	34.10	20.20
AEZ range classification	312–323	311–324
Thermal climate (AEZ classification)	Tropics	Tropics

Table 2. Amount of samples per set of database

Database	Nigeria	Ethiopia	Total
Training	606	524	1130
Testing 1	594	524	1118
Testing 2	582	521	1103
Total	1782	1569	3351

Agriculture Organization of the United Nations (FAO) created the Global Agro-Ecological Zones Methodology (AEZ) [14] to classify agricultural potential taking into account agro-climatic, soil and terrain conditions. Both, Nigeria and Ethiopia are located in the same range of AEZ classification, having as a common characteristic their thermal climate.

1.2. Data management

To develop this methodology, the household-level information obtained for each country was divided in three sets of databases: training (used to formulate potential MLR models for each appliance), testing 1 (used for testing the potential models in order to select the best-fitting one) and testing 2 (used to test the selected model). The details of these sets are presented in Table 2. Each of them was chosen randomly, taking into account rural households from all the states; consequently, these sets were aggregated to a state and national level to test the methodology at different spatial scales. The data were cleaned from outliers and organised considering more than 10 variables related to demographic, socio-economic, climate and geographic aspects.

2. Methodology

The MLR method was selected for the estimation of appliances ownership, using the training database with household-level information to find the correlation between each appliance ownership and multiple independent variables. The driver variables were found by applying an Analysis of Variation (ANOVA) F-static test. These variables are divided in household-level (size, proximity to major road, and electricity grid access), and head of household (age, literacy, religion, and salaried employment situation) (see section 2.1). Most appliances diffusion methodologies (used to project the diffusion of every appliance over time at a national level), apply logistic regression models for a binary choice projection, giving results ranging from 0 to 1 – where 1 represents the complete diffusion [8,9].

The objective of this paper is to propose a methodology to estimate the current appliances ownership, defined in this context by the number of appliances owned per rural household. For example, in the case of the appliances chosen for this study, one household can have more than one mobile phone or radio; therefore, it is possible to obtain a value larger than one for the appliances ownership. The MLR function is given by (1):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon, \tag{1}$$

where y is the appliances ownership, β_0 is the intercept (the value of y when all the variables are 0), $\beta_{1\ldots n}$ are the regression coefficients, $x_{1\ldots n}$ represent the independent variables, and ε is the error term.

2.1. Driver variables

The variables contained in the databases were classified as:

- Demographic: household size, sex and age of household head;
- Socioeconomic: aggregated annual income and electricity access of the household, employment situation, literacy, years of education, and religion of household head;
- Geographic: proximity to major road, proximity to population centre;
- Climate: average annual temperature and precipitation

The correlation between all these variables was first tested in order to create a set of independent variables to be used in the model for each appliance. The selection of the independent variables was done by applying an Analysis of Variation (ANOVA) F-static test. This test represents the ratio of the variation between means of the sample groups (external variance) and the variation within the samples (internal variance) [16], as described in (2).

$$F = \frac{\sum_{i=0}^{k} n_0 \left\| \bar{X}_{i.} - \bar{X}_{..} \right\|^2 \big/ (k-1)}{\sum_{i,j} \left\| X_{i,j} - \bar{X}_{i.} \right\|^2 \big/ (n-k)}, \tag{2}$$

$$\bar{X}_{i.} = \sum_{j=1}^{n_i} \frac{X_{ij}}{n_i}, \quad \bar{X}_{..} = \sum_{i=1}^{k} \frac{n_i \bar{X}_{i.}}{n}, \quad n = \sum_{i=1}^{k} n_i, \tag{3}$$

In (3), $\bar{X}_{i.}$ represents the sample mean in the i-th group, n_i the number of observations in the i-th group,

Table 3. Description of variables at household level

Variable	Description	Type	Unit or range
Size	Number of people living in the household	Continuous	People
Proximity to major road	Distance between the household and the closest major road	Continuous	km
Electricity access	Only considering grid-connection	Discrete	0–1[1]

[1]one indicating grid-connection, zero otherwise.

Table 4. Description of head of household variables

Variable	Description	Type	Unit or range
Age	Age of the household head	Continuous	years
Literacy	The household head can read and write in at least one language	Discrete	0–1[1]
Religion	The household head has a religion	Discrete	0–1[1]
Salaried employment	The household head has been employed with a fixed salary over the last year	Discrete	0–1[1,2]

[1]one indicating a positive answer, zero otherwise.
[2]without considering agriculture or home-productivity activities.

the overall mean of the data, X_{ij} is the j-th observation in the i-th out of k groups, and n is the overall sample size. MLR models are used to test the individual effects of many explanatory variables on a response; and the collinearity of these variables often threatens their statistical interpretation [16]. It was found that a strong correlation exists between some of the variables selected for this methodology; for example, the household head literacy level is positively correlated with the electricity access of the household. By applying a Global F test for each of the MLR models of the studied appliances, the individual effect of the considered variables on estimating the ownership of each appliance was tested. The selected variables are included in the Tables 3–4.

2.2. Model accuracy

The accuracy of the models for each appliance was measured by applying the root mean square error (RMSE). This method was chosen because the error of most of the models present a Gaussian distribution centred at zero. It provides a complete picture of the error distribution of the model when compared to the real values; also, it is considered as more reliable in statistics when having large sample sizes [17, 18].

Table 5. Comparison of values of R^2 for each appliance's model using different training sets

Appliance	Nigeria	Ethiopia	Combined
Television	0.40	0.78	0.54
Radio	0.06	0.02	0.37
Mobile Phone	0.24	0.40	0.22
Refrigerator	0.27	0.31	0.34

3. Results and discussion

In this section, the results from the appliances ownership model for each of the analysed appliances and the error distribution when the model is tested in the different aggregation sets (household, state and national level) are discussed. Finally, the GIS application is presented showing different examples of the implementation of this approach to identify potential locations of future electrification projects.

3.1. Appliances ownership modelling

The four appliances considered for this study are described in more detail below. To find the best-fitting model for each appliance, three different training sets were tested: one including training data from each individual country, and one combining both countries' data. This process was done by comparing the fitting performance of the models measured by the R^2 value. As presented in the Table 5, using the training data from Ethiopia gave a greater value of R^2 in certain cases, however, the combined model was selected because the large sample size produces less uncertainty and more statistically meaningful results [17]. The parameters of the final models are presented in Tables 6–9.

3.1.1. Televisions

Previous studies have found that televisions are by far the most desirable electrical appliance among rural households, after lighting [8]. This appliance tends to increase its economic accessibility with the time, which is demonstrated in this study by obtaining a low correlation between the ownership of this appliance and the household's income. As expected, televisions ownership presents a high dependency on the electricity

Table 6. MLR model results for televisions

R^2	Residual Std. Error	Intercept (β0)	Size (β1)	Electricity Access (β2)	Literacy (β3)
0.54	0.26	–0.33	0.03	0.67	0.43

Table 7. MLR model results for radios

R^2	Residual Std. Error	Intercept (β0)	Size (β1)	Age (β2)	Religion (β3)
0.37	0.36	0.55	0.02	0.002	–0.50

Table 8. MLR model results for mobile phones

R^2	Residual Std. Error	Intercept (β0)	Proximity to Major Road (β1)	Age (β2)	Salaried Employment (β3)
0.22	0.36	–0.16	–0.01	0.02	1.06

Table 9. MLR model results for refrigerators

R^2	Residual Std. Error	Intercept (β0)	Size (β1)	Electricity Access (β2)	Literacy (β3)
0.34	0.14	–0.28	0.04	0.16	0.16

access of the household, showing also that households connected to the grid tend to have more than one television. Another variable that affects the ownership modelling is the household head's literacy, which is positively correlated to the electricity access. The household size is also considered in the model due to its significance on the model's performance, even if it has less impact than the other two variables.

3.1.2. Radios

According to household surveys, radios are among the electrical appliances mostly owned in rural areas of developing countries mainly due to its low cost. In this study, it was found that the household size, the age and religion of the household head has an influence on the ownership of this appliance. The household tends to have more radios, if it has a large size, if the head is older and if he is not a religious person. It is important to highlight how having a religion in the rural areas of both countries has a negative correlation with the household size, the education level and employment situation of the household head. It was also found that the households located at a larger distance from the nearest population centre (considering as population a site inhabited by more than 20,000 people, according to [10]) and major road, tend more to have a religion. These isolated sites have limited access to signal transmission towers, so the radios can barely work or not at all, consequently, the radios ownership is lower.

3.1.3. Mobile Phones

Mobile phone usage has increased dramatically in developing countries over the last years independently of their electrification rate [19]. In this analysis, it was found that a positive correlation exists between the ownership of mobile phones and the household head employment situation (if head of the family has a salaried employment) and age, while a negative one for the household's distance to the closest major road. Generally, the telecom transmission towers are located near the major roads and large population centres; therefore, if the household is isolated from these, having a mobile phone will not be functional.

3.1.4. Refrigerators

Refrigerators are considered as highly wanted but relatively expensive for rural households [8]; also, this appliance is considered as one of the most energy consuming in the global residential sector. However, the annual income of the household had a relatively low influence on its ownership and the overall performance of the model; therefore, it was decided not to include it. In this analysis, the variables evaluated in the model for televisions were also suitable to project the ownership of refrigerators.

The models presented above were chosen based on the highest value of R^2, even though if these were below 0.50 (with exception of the televisions model). In order to discard the possibility that the low R^2 values were obtained due to the limited capacity of the MLR method on estimating the appliances ownership, other linear and non-linear regression methods were tested to analyse their performance on fitting the training dataset. Using the model for televisions as example, the comparison of these methods based on their R^2 value is presented in Figure 1. The methods used to compare the MLR's results are the following:

- Bootstrap Aggregating (BAGGING): generates multiple versions of a predictor using them to create aggregated averages to estimate a numerical outcome [20];
- Support Vector Machine (SVM): finds a non-linear function that has at most a certain deviation from the obtained targets for the training data [21];
- Generalised Linear Models (GLM): includes iterative weighted linear, logistic and Poisson regression to obtain maximum likelihood estimates of the parameters [22];
- Generalised Boosted Regression Modelling (GBM): creates models in a sequence based on the training data attempting to correct deviations from the previous model until the best prediction is made [23].

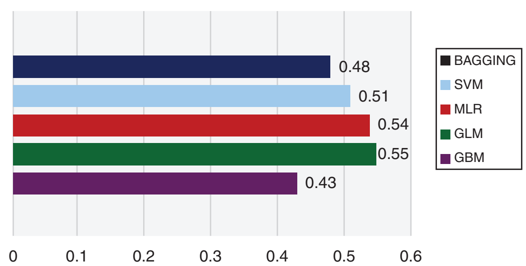

Figure 1: Sensitivity analysis of fitting performance of different methods based on R^2 values for the televisions model

In the Figure 1 is shown that the fitting performance has almost no difference when applying the different methods, while the GLM and MLR are giving as result the highest values of R^2. This demonstrates the difficulty of finding clear drivers of the acquisition of these appliances at a household level, mainly for the cases of small appliances such as radios or mobile phones which can commonly be owned more than once in a household. This also explains why many studies are based on more aggregated data (state or national level) and are not focused on determining the appliances acquisition at a household level, which is mainly a function determined by human behaviour rather than evident drivers to find acquisition patterns.

3.2. Error distribution

The error distribution for each model tested at a household and state level is presented in the Figures 2 and 3. As mentioned before, most of the models have a Gaussian distribution centred at zero, which is a good indicator that the models generally produce accurate results without bias towards an under or overestimation

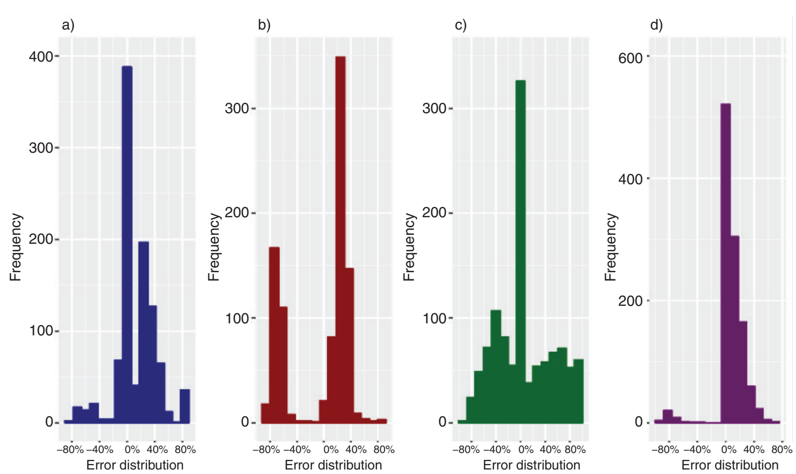

Figure 2: Error distribution of models tested at household level, a) televisions, b) radios, c) mobile phones, d) refrigerators. Number of observations: 1,118

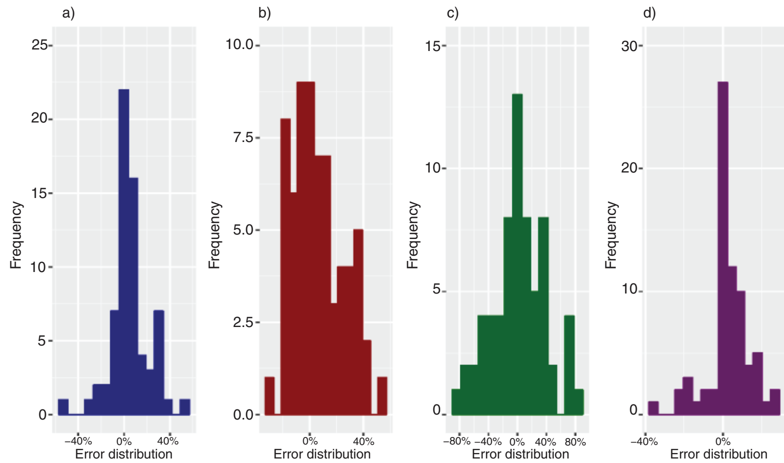

Figure 3: Error distribution of models tested at state level, a) televisions, b) radios, c) mobile phones, d) refrigerators. Number of observations: 66.

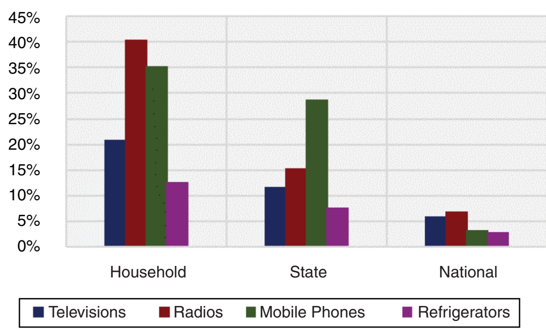

Figure 4: Errors estimated with RMSE for the selected appliances ownership models at every spatial scale.

of the ownership for the appliances studied. Exemptions were found in the cases of the household level test for radios (in which most of the ownership was projected within an overestimation range of 11-35%), and in the state level test of the refrigerators (having more than 20 results with an error between 0-9%). This is attributed to the existence of outliers and lack of representativeness in the database. In Nigeria the probability of rural households owning at least one radio and refrigerator is of 76% and 9% respectively; while in Ethiopia is only of 23% and 2%, respectively. Therefore, including Nigeria in the models' training databases sets a high expectation for these appliances ownership values, which leads to a small overestimation.

As expected, the household level tests presented the least accuracy (measured by the RMSE); this is attributed to the differences in human behaviour and particular influence of each household. The accuracy of the models increased together with the spatial aggregation of the datasets (see Figure 4). For the state and national level, the models presented a lower error, which proves that all the outliers from the household level were levelled out in the aggregation process. The television model applied at a household level had an error of 20.9%, at a state level of 11.6% and at a national level of 6%. In the case of the radios, the errors were of 40.5%, 15.3% and 7% respectively. The mobile phones models gave errors of 35.3%, 28.7% and 3.2% respectively, while the refrigerators presented the lowest errors at all levels with values of 12.6%, 7.7% and 2.8%. It is important to note that the errors obtained at a household scale might have an implication on the estimation of energy demand for distributed power generation projects at a community scale.

The models for the ownership of televisions and refrigerators presented a high correlation with the electrification rate, which was expected due to the high electrical consumption that these appliances require. Taking as an example the televisions model applied at a state level, in Figure 5.a, the correlation can be observed for the case of Nigeria (with a R^2 of 0.88), which presents the highest ownership value and electrification rate. In the case of Ethiopia, the dependency is very low, obtaining a R^2 of 0.03. Following the example of Nigeria, Figure 5.b gives a representation of the difference between the modelled ownership of televisions and the real values per observation, which demonstrates the accuracy of the estimation.

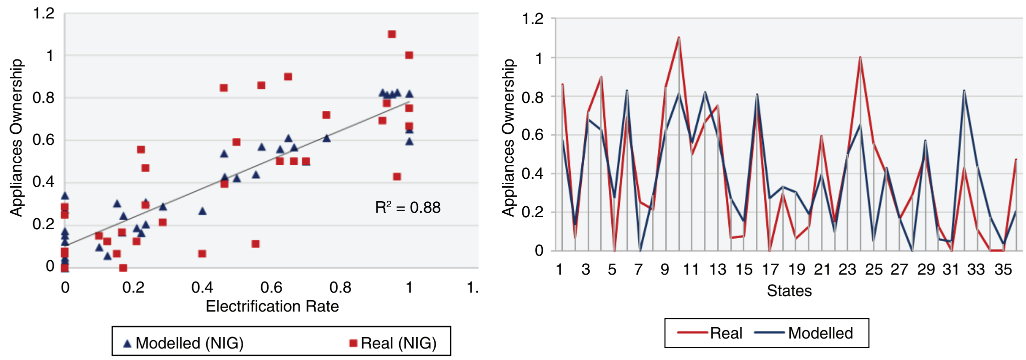

Figure 5: Example of Nigeria: a) Correlation between televisions ownership and electrification rate. b) Difference between modelled and real televisions ownership rate values

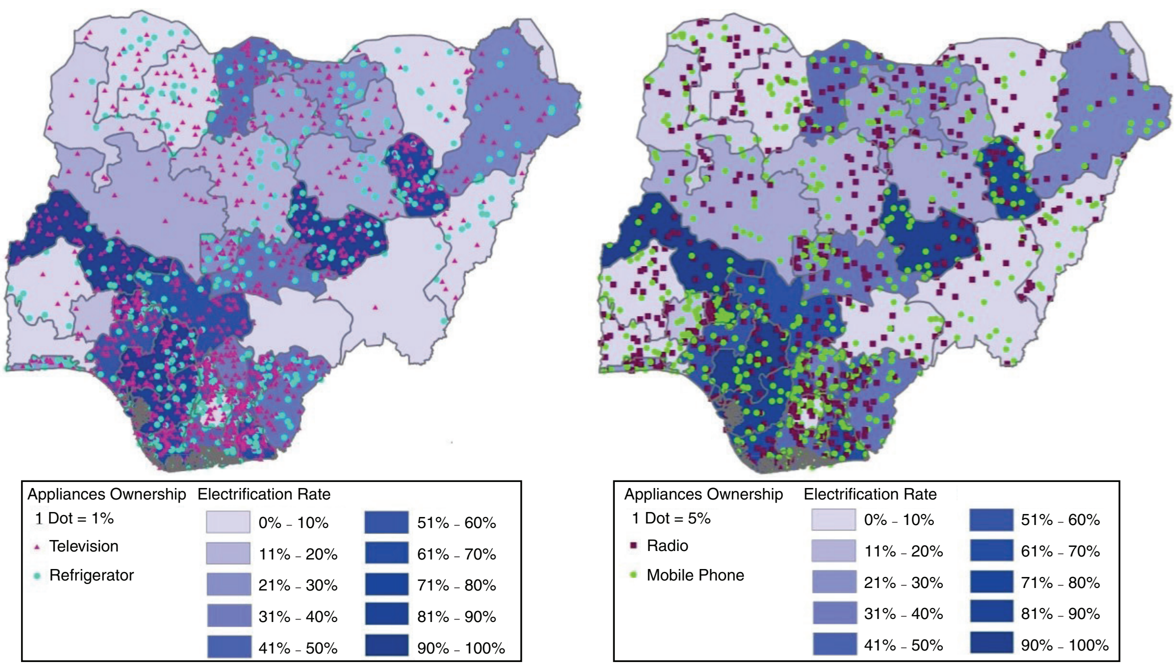

Figure 6: Representation of the ownership for the modelled appliances for Nigeria (left figure: televisions and refrigerators, right figure: radios and mobile phones) together with the rural electrification rates per state. The dots represent the percentage (stated in the scale) of ownership per household

3.3. GIS application

The presented models can be used to geographically indicate in which places of a country the highest electrical appliances ownership values are located, and to analyse if these values are associated to the electrification rates in those places. This application can be very useful for the national authorities to identify potential sites for future electrification projects. Figures 6 and 7 present the modelled ownership for each appliance showed by country and state, as well as the rural electrification rates for the year 2015-2016 of the country.

In both countries, it is observed that the ownership for televisions and refrigerators are concentrated in the regions with the highest electrification rates. However, in the case of the mobile phones and radios, the ownership values are more disperse and they are not necessarily located in places where the electrification rates are higher. This is attributed to the fact that these two small appliances are economically accessible to almost anyone and they do not consume a high amount of electricity. Evidence has shown that people in rural areas are likely to use dry-cell/wet-cell batteries to

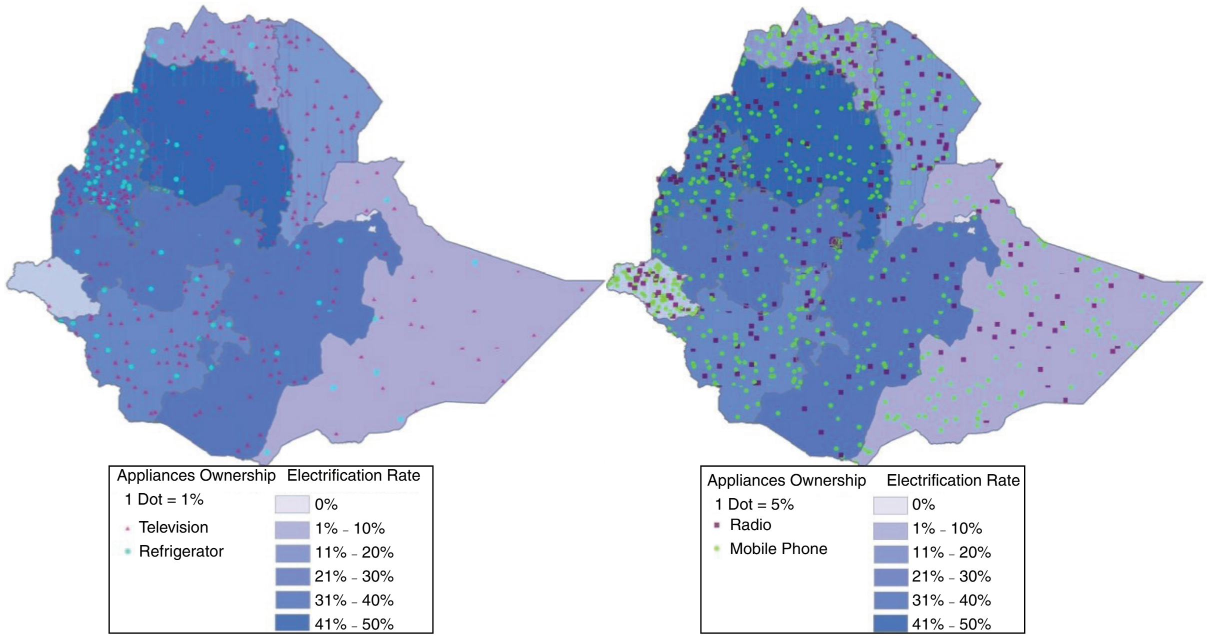

Figure 7: Representation of the ownership for the modelled appliances for Ethiopia (left figure: televisions and refrigerators, right figure: radios and mobile phones) together with the rural electrification rates per state. The dots represent the percentage (stated in the scale) of ownership per household

power these electronic devices [24]. It is important to mention that the number of observations in the state of Imo and Oyo (Nigeria) were low, and the rural population surveyed had from zero to 10% of electrification, nevertheless these states are the ones with the highest electrification rates in the country, having 70% and 67% respectively [26].

In Figure 8, the national transmission and distribution grid is projected against the appliances ownership per state for both countries. Both national grids data were obtained from [27], in the case of Ethiopia, it was collected in 2014, while for Nigeria in 2015. *Planned* transmission and distribution grids refers to these that are planned to be delivered by the years 2020–2025. Potential states could be identified to be listed as priority for the development of future electrification projects. For Nigeria, it is observed that in the states with the lowest ownership of high power-consuming appliances (televisions and refrigerators), the infrastructure of the national grid is poor (as it happens in the states of Taraba and Oyo, for example). There are exceptions, such as the case of Borno, which has a relatively high ownership of televisions and refrigerators but the national grid is not well developed in this area yet. The states mentioned have rural electrification rates ranging from 10–30% (with the exception of Oyo). In the case of Cross River and Katsina, some areas of these states have a good connection to the national grid; however, almost 50% of their rural population is not connected. Moreover, it is shown that these people own a relatively high amount of televisions and refrigerators. For Ethiopia, it can be easily noticed that the state of Somali lacks of electricity network infrastructure (having up to 10% of rural grid connection), while it is presented that these households own large appliances (televisions and refrigerators). This is also the case for the state of Benshangui-Gumaz, which shows the largest television ownership with 38%, and refrigerators with 18%, while still approximately 64% of the rural households do not have a grid connection.

The presented geographic information can give an insight of which electrification solution is the most appropriate for each of the highlighted states (considered as potential priority) from the countries analysed in this work.

To give a representation of the number of appliances owned per state, the amount of rural households is estimated for Cross River in Nigeria (Figure 8). This

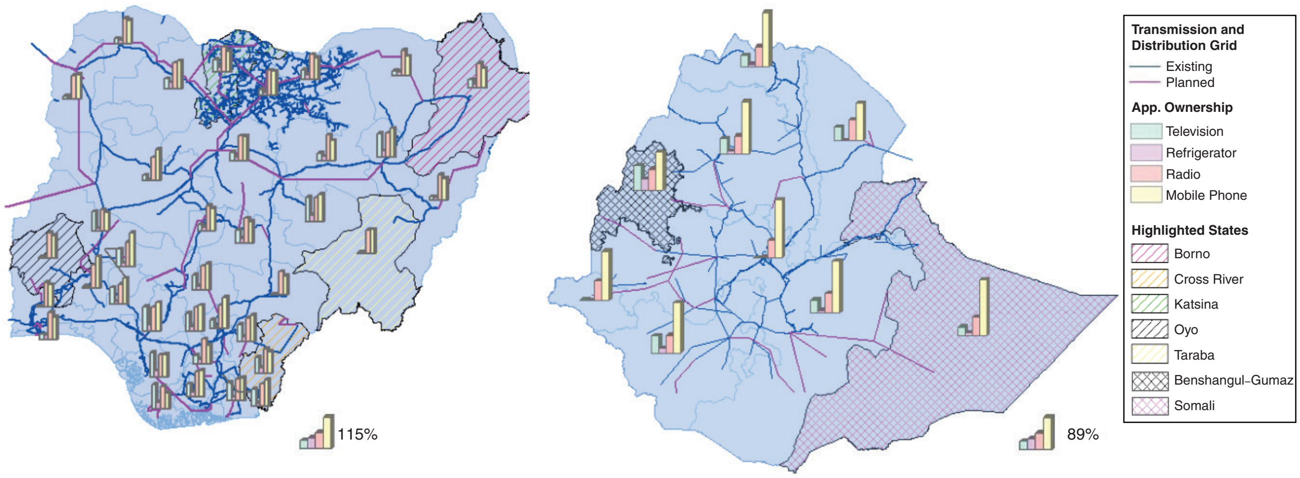

Figure 8: Projection of the national transmission and distribution grid (existing and planned) and the appliances ownership per state (left figure: Nigeria, right figure: Ethiopia). The number next to the bar scale represents the value of the longest bar in the chart for each country

Table 10. Estimation of the number of appliances for the state of Cross River, Nigeria

Rural population (2015) [10]	Average HH size in rural areas [27]	Number of HH	Televisions		Refrigerators		Radios		Mobile Phones	
			App. Own.	Amount	App. Own.	Amount	App. Own.	Amount	App. Own.	Amount
2,865,429	5.10	561,848.82	0.54	303,398	0.2	112,370	0.83	466,335	0.84	471,953

amount of households is then multiplied by the calculated ownership of each appliance. These estimations are presented in Table 10.

4. Conclusion

Rural electrification is a global priority for policymakers as it contributes significantly to the development of a country. The methodology presented in this paper gives a sight of the potential of using existing and widely available data from rural households' surveys for societal benefits, such as supporting the planning of electrification projects. It estimates the number of electrical appliances owned by a rural household according to its socioeconomic, demographic and geographic characteristics – which can be used to calculate the electricity demand per household for a certain period, by considering the power rates and time of use (e.g. daily or yearly) of these appliances. The model for each studied appliance presented a relatively low error due to the amount of samples considered to create the linear regressions. The multiple coefficient of determination (R^2) was low for most models; this demonstrates the difficulty on determining acquisition patterns at household level using evident drivers. In addition, it can be attributed to the existence of outliers in the databases, even if more than 1000 samples were removed from each country due to data discrepancies. The model for refrigerators was the most accurate, having errors ranging from 12.6% at a household level and 2.8% at a national level. On the other hand, the models for the radios and mobile phones had the highest errors mainly at a household level with 40.5% and 35.3% respectively. Due to the high ownership rates of these appliances and their low correlation with the evident and well-known socioeconomic, demographic and geographic aspects, the precise projection of the ownership at a household level becomes a difficult task. In future research, the methodology will be improved by adding country scale variables to not only consider a bottom-up approach, analysing the changes in the models' fit. It will be tested in other countries from sub-Saharan Africa with similar socioeconomic and geographic characteristics in order to reach its regional parameterisation, and then their current and future electricity demand will be estimated based on the expected future changes on acquisition patterns. In addition, the

GIS application will be improved to give accurate recommendations of specific sites were potential electrification projects can be developed.

Acknowledgments

This research has been financially supported by the ETH Grant Progboram, under the project: *"Rural energy demand modelling in developing countries"*.

References

[1] International Energy Agency. World Energy Outlook; 2017. http://www.iea.org/weo2017/

[2] Kirubi, C, AJ, Community-based electric micro-grids can contribute to rural development: evidence from Kenya. World Development 37(7) (2009) pages 1208–1221. https://doi.org/10.1016/j.worlddev.2008.11.005

[3] Deshmukh, R, JP, Sustainable development of renewable rnergy mini-grids for energy access: a framework for policy design. Berkeley: Lawrence Berkeley National Laboratory; 2013. http://eta-publications.lbl.gov/sites/default/files/lbnl-6222e.pdf

[4] Howells, MI, TD, A model of household energy services in a low-income rural African village, Energy Policy 33(14) (2005) pages 1833-1851. https://doi.org/10.1016/j.enpol.2004.02.019

[5] Mandelli, S, JB, Off-grid systems for rural electrification in developing countries: Definitions, classification and a comprehensive literature review, Renewable and Sustainable Energy Reviews 58 (2016) pages 1621–1646. https://doi.org/10.1016/j.rser.2015.12.338

[6] Kwakwa, P, Adu, G, Osei-Fosu, A, A time series analysis of fossil fuel consumption in Sub-Saharan Africa: evidence from Ghana, Kenya and South Africa. International Journal of Sustainable Energy Planning and Management 17 (2018) pages 31–44. dx.doi.org/10.5278/ijsepm.2018.17.4

[7] Kadraoui, H, Benouaz, T, Bekkouche, S, Analysis of energy consumption for Algerian building in extreme North-African climates. International Journal of Sustainable Energy Planning and Management 19 (2019) pages 45–58. https://doi.org/10.5278/ijsepm.2019.19.5

[8] McNeil, MA, VE, Modeling diffusion of electrical appliances in the residential sector, Energy and Buildings 42(6) (2010) pages 783–790. https://doi.org/10.1016/j.enbuild.2009.11.015

[9] McNeil, MA, VE, Forecasting electricity demand in developing countries: a study of household income and appliance ownership. Berkeley: Lawrence Berkeley National Laboratory; 2014. https://www.researchgate.net/publication/228619092_Forecasting_Electricity_Demand_in_Developing_Countries_A_Study_of_Household_Income_and_Appliance_Ownership

[10] Narasimha, D, Ummel, K, White goods for white people? Drivers of electric appliance growth in emerging economies, Energy Research & Social Science 27 (2017) pages 106–116. https://doi.org/10.1016/j.erss.2017.03.005

[11] World Bank. Living Standards Measurement Study; 2017. http://econ.worldbank.org/>

[12] International Fund for Agricultural Development; 2017. https://www.ifad.org/who/overview

[13] Williams, NJ, PJ, Enabling private sector investment in microgrid-based rural electrification in developing countries: A review, Renewable and Sustainable Energy Reviews 52 (2015) pages 1268–1281. https://doi.org/10.1016/j.rser.2015.07.153

[14] IIASA/FAO. Global Agro-Ecological Zones (GAEZ v3.0). Laxenburg, Rome: International Institute for Applied Systems Analysis (IIASA), Food and Agriculture Organization of the United Nations (FAO); 2017. http://www.fao.org/nr/gaez/en/

[15] World Bank. Country Indicators Database; 2017. https://data.worldbank.org/indicator/

[16] Cuevas, A, MF, An anova test for functional data, Computational Statistics & Data Analysis 47(1) (2003) pages 111–122. https://doi.org/10.1016/j.csda.2003.10.021

[17] Williams, R. Review of Multiple Regression, South Bend: University of Notre Dame; 2015. https://www3.nd.edu/~rwilliam/stats2/l02.pdf

[18] Chai, T, Draxler, RR, Root mean square error (RMSE) or mean absolute error (MAE)? - Arguments against avoiding RMSE in the literature, Geoscientific Model Development 7(3) (2014) pages 1247-1250. https://www.geosci-model-dev.net/7/1247/2014/gmd-7-1247-2014.pdf

[19] Martinez-Cesena, EA, PM, Using mobile phone data for rural electrification planning in developing countries. Manchester: University of Manchester; 2015. https://arxiv.org/abs/1504.03899

[20] Breiman, L, Bagging Predictors, Machine Learning 24(2) (1996) pages 123-140. https://doi.org/10.1007/BF00058655

[21] Smola, A,J, A tutorial on support vector regression, Statistics and Computing 14(3) (2004) pages 199–222. https://doi.org/10.1023/B:STCO.0000035301.49549.88

[22] Nelder, J,A, Generalized Linear Models, Royal Statistical Society 135(3) pages 370-384. https://doi.org/10.2307/2344614

[23] Friedman, J, Greedy Function Approximation: a Gradient Boosting Machine, The Annals of Statistics 29(5) pages 1189–1232. https://statweb.stanford.edu/~jhf/ftp/trebst.pdf

[24] Adeoti, O, BO, Solar photovoltaic-based home electrification system for rural development in Nigeria: domestic load

assessment, Renewable Energy 24(1) (2011) pages 155–161. https://doi.org/10.1016/S0960-1481(00)00188-9

[25] DIVA-GIS. Country Level GIS Data; 2017. http://www.diva-gis.org/Data

[26] World Bank. International Development Association. Nigeria Electrification Project; 2018. http://projects.worldbank.org/P106172/nigeria-electricity-gas-improvement-project-negip?lang=en&tab=overview

[27] World Bank. Africa Electricity Grids Explorer; 2018. http://africagrid.energydata.info/

[28] USAID. Demographic and Health Surveys; 2017. https://dhsprogram.com/