ISSN 2004-2965 Energy Proceedings, Vol. 20, 2021

International Conference on Applied Energy 2021
Nov. 29 - Dec. 2, 2021, in Bangkok, Thailand
Paper ID: 578

Understanding the Multidimensionality of Energy Access in Remote Areas of Developing Countries Using a Multivariable Statistics Approach

Alonso Alegre-Bravo[1], Lindsay Anderson[1]

[1] Cornell University- Department of Biological and Environmental Engineering- Ithaca, NY- United States

ABSTRACT

The COVID-19 pandemic has negatively affected people living in vulnerable conditions around the world—especially rural communities with lack of access to basic services. Rural electrification can play a fundamental role in pandemic recovery by facilitating the provision of basic services and improving quality of life for poor rural communities. This study aims analyze this potential through understanding the correlation of multidimensional variables with electricity consumption. The dominant approach to rural electrification has been a for-profit model, which has hindered electricity access for the most vulnerable populations of the global South. Using multiple linear regression tools, this study confirms that a renewable energy generation model that considers social, economic and environmental factors, in addition to technical factors, is pivotal to increasing electricity consumption, and consequently facilitate greater energy access in developing countries.

Keywords: energy access, electricity consumption growth, Global South, renewable energy, multivariable analysis, multiple linear regression

NONMENCLATURE

Abbreviations	
GDP	Gross Domestic Product
HDI	Human Development Index
LASSO	Least Absolute Shrinkage and Selection Operator
OLS	Ordinary Least Squares
SDG	Sustainable Development Goals
UN	United Nations
VIF	Variance Inflation Factor

1. INTRODUCTION

In 2021, 10% of the world's population still suffers from lack of access to electricity [1]. The COVID-19 pandemic has exacerbated energy poverty around the world by eliminating some development achievements accomplished in the last decade and creating new barriers to accessing electricity [2]. Lack of or poor electricity access disproportionately affects rural communities. Resolving this inequality is one of the Sustainable Development Goals (SDG) established by the United Nations (UN). One challenge faced in rural electrification planning is determining the electricity consumption growth in communities. In rural contexts, it is common to find electrification projects with over-consumption, meaning communities end up using more electricity that the project had been designed for [3]. This situation jeopardizes the electrical systems operation, and results in beneficiaries being dissatisfied with the project. In addition, energy utilities, which oversee rural electrification programs, base their decisions on standards or criteria that prioritize cheap projects and for-profit models [4]. To avoid these problems, rural electrification developers need to recognize that energy access is a multidimensional challenge and design projects with more diverse inputs in addition to the technical data. By including social, economic, demographic, and environmental data, rural electrification developments can achieve a more extended lifespan.

Energy access researchers have taken a variety of approaches to understanding the energy poverty problem. Researchers have used three levels of assessment: macro, meso and micro level, and using methodologies such as multi-criteria decision-making tools, regression analysis, life cycle analysis, multiple linear regression, uncertainty and sensitivity analysis [5]. Most of the literature in energy poverty, energy access or rural electrification applies a micro level analysis. A micro level assessment examines indicators such as: housing characteristics, energy efficiency of technologies used, socio-economic and demographic household characteristics, health indicators, or household income. Authors like [6], [7] and [8] examine energy poverty under a micro level assessment using multiple linear regressions. Their analysis includes comparisons between the efficiency performance of small-scale renewable energy technologies, solar photovoltaic, and socio-economic indicators in education, health and income. In contrast, energy poverty research using macro level assessment is not that common. Macro level assessment zooms out to view the entire problem and identifies the most significant challenges in combating energy poverty. Macro level assessments evaluate indicators such as greenhouse gas emissions, environmental energy policy actions, energy prices, energy consumption, energy balances, and social policy actions. This study uses a macro level assessment and a multiple linear regression methodology.

To have a broader understanding of the energy poverty problem, the objective of this study is to analyze four socio-economic indicators, one environmental and one technical indicator to examine their behavior and correlation in terms of electricity consumption using a multivariable data analysis. This study works with three rural regions of Latin America: Alta Verapaz in Guatemala, Jinotega in Nicaragua, and Huánuco in Peru. These regions share two common characteristics for developing rural electrification projects. First, they have a low electrification access rate [9-11]. Second, they are in the mountains, having direct access to sources of water; in the energy context, this means that small-hydro power projects could be potentially developed. This study focuses on small-hydro technology because among the small-scale power sources, it is a substantially cost-effective technology in the long term in comparison to other conventional technologies [12].

From a technical approach, the study analyses the electricity consumption rates per capita (kWh/year) of the countries of interest and renewable energy generation (%). The demographic and socio-economic data examine: i) population growth (%); ii) GDP per capita (US dollars); iii) rural population (%); iv) school enrollment, tertiary level (%) and the environmental data examines v) carbon dioxide emissions per capita (metric tons per capita) [13].

Data for these specific three regions in Latin America is being collected and to date is not yet available. Therefore, this study leverages data from the macro-level, examining national indicators instead of specific regional indicators, to have a baseline for understanding the relationship between these indicators and electricity consumption growth. The countries are chosen based on their similar Human Development Indices (HDI) to the rural regions in Guatemala, Nicaragua and Peru. This index consists of three major dimensions, long and healthy life, knowledge, and a decent standard of living [14].

2. METHODS

This study examines technical, environmental and socio-economic data using a multivariable linear regression. The aim is to run a multivariable linear regression and identify (1) whether variables are correlated, and (2) the statistical significance of the correlation of such variables to the dependent variable, electricity consumption growth. The basic form of the proposed model used in the regression is:

$$EC = \beta_0 + \beta_1 PG + \beta_2 GDP + \beta_3 RP + \beta_4 SE + \beta_5 CO2 + \beta_6 RE \quad (1)$$

In equation 1, the dependent variable *EC* (electricity consumption per capita) is the value in kWh per capita of electricity consumption per year in each country. The electricity consumption per capita was obtained dividing the total electricity consumption (billion kWh) by the total population of each country. The independent variables are: i) *PG* (population growth %), ideally, the greater the population the higher the energy consumption will be, providing a positive relationship with *EC*; ii) *GDP* (gross domestic product per capita) represents, in theory, the income per person per year. This varies between countries depending on their economic inequality. In this study, it is expected that the greater the income the higher the quality of life and EC; iii) *RP* (rural population) is the percentage of the country population living in remote areas with more limited basic services relative to urban areas. Because rural populations tend to be socioeconomically marginal and have lower incomes, this study expects that countries with higher rural populations will have lower energy consumption rates; iv) *SE* (school enrollment, tertiary) is the percentage of the population that has satisfactorily concluded high school and is enrolled in a university level. The expectation is that the greater the *SE* the higher *EC* will be [15]. Ideally, a society with more access to education has opportunities to achieve higher salaries and therefore consume more electricity. In the environmental context, this study analyzes v) *CO2* (carbon emissions per capita). Due to the diversity of sources for electricity generation and energy use in the Global South is not clear if higher the CO_2 emissions per capita will correlate to higher *EC.* In the technical context, this study assesses vi) *RE* (% of renewable energy consumption). This study examines *RE* because in 2021 renewable energy is an important vehicle to reduce the emissions of CO_2 into the atmosphere and combat climate change. The UN, through the Intergovernmental Panel on Climate Change, urges governments to invest in clean energy sources instead of fossil fuels [16, 17]. At this stage of the study, it is not clear if higher the *RE* investment will correlate to higher *EC*.

The selection of these variables was influenced by availability of data for the most recent years. The World Bank, the U.S. Energy Information Administration and the International Energy Agency possess all the data for these variables for the year 2018 [18-25]. To organize the data, several countries that have similar HDI to remote areas in Guatemala, Nicaragua, and Peru were examined. In the case of the studied regions in Guatemala, Nicaragua, and Peru, their HDI fit in the Medium Human Development and Low Human Development groups [9-11]. The HDI range varies between 0.7 and 0.39, encompassing 75 countries, mostly from Africa, Latin America, and South Asia.

3. THEORY

Having several variables with different units required the standardization of values. Equation 2 depicts the typical standardizing equation [26]:

$$Z = \frac{x - \mu}{\sigma} \qquad (2)$$

Where Z is the normalized score, x the observed value, μ the mean of the sample, and σ the standard deviation. With standardized data an ordinary least squares (OLS) regression is used to determine which values are more correlated to electricity consumption per capita. A prediction model is used to calculate the root mean squared error (RMSE). To corroborate the correlation of covariances, the study calculates a variance inflation factor (VIF) to quantify the multicollinearity in the OLS and identify inflation of the standard error [27].

To decrease complexity and prevent over-fitting from OLS, this study implements Ridge and LASSO regression to penalize the function [28]. After calculating OLS, Ridge and LASSO fitted regression coefficients, the study compares coefficient values of the independent variables. Energy access literature using multiple linear regression mostly implement OLS analyses. Approaches to reduce multicollinearity like Ridge and LASSO are not very common. This study explores LASSO and Ridge approaches to show a different perspective of multivariable analysis in energy access literature.

4. RESULTS

The OLS results show that r squared has a value of 0.7716, meaning that approximately 77% of *EC* can be explained by our model. Table 1 depicts the behavior of independent variables in terms of *EC*.

Table 1. Coefficients OLS Regression

	Estimate	Standard Error	t value	Pr (>|t|)
RE	0.15665	0.08514	1.840	0.07870*
PG	0.07512	0.09474	0.793	0.43598
GDP	0.45799	0.24309	1.884	0.07226*
RP	0.19350	0.13782	1.404	0.17366
SE	0.15577	0.11106	1.403	0.17412
CO2	0.36206	0.10067	3.597	0.00152**

According to the OLS regression, the only three predictors significant at the 10% level are *RE, GDP* and *CO2*. *CO2* is the only predictor significant at the 1% level. This result is compared in the next regressions with LASSO and Ridge. Table 1 also displays that *EC* increases when also *PG*, *GDP*, *RP*, *SE* and *CO2* are higher. The root mean square error (RMSE) is calculated to understand how spread the data is around the line of best fit. The RMSE in this study is 0.3383, as Figure 1 shows the data for the most part is concentrated. An explanation for obtaining this RMSE could be the number of independent variables. This spread of data would impact the outcome of the analysis since it would be more predictable to make a theoretical confirmation.

The analysis of the variance inflation factor (VIF) for each covariate identified that the largest VIF is *GDP* with a value of 6.2719. Having small VIFs indicate that the values have low correlation among independent variables and confirms that redundant information does not exist among them.

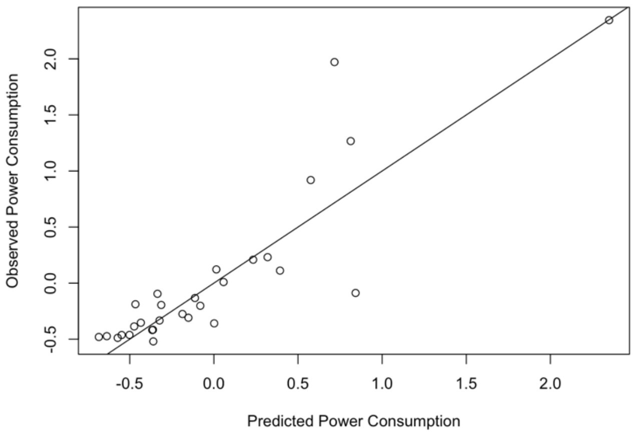

Figure 1. Predicted Power Consumption vs. Observed Power Consumption

To reduce collinearity in the linear model analysis and allocated weights more precisely, LASSO and Ridge regressions are also used in this analysis. LASSO and Ridge compares the mean cross-validated error values in terms of lambda values. Table 2 presents a fitted regression coefficient for the OLS, LASSO and Ridge regression models using the studied covariates.

Table 2. OLS, LASSO and Ridge Regressions and Covariates Fitted Regression Coefficients

	OLS	LASSO	Ridge
RE	0.157	0.892	0.885
PG	0.075	0.000	-0.031
GDP	0.458	0.000	-0.039
RP	0.194	0.000	0.021
SE	0.156	0.000	-0.003
CO2	0.362	0.000	-0.004

In Table 2, contrary to OLS regression, LASSO allocates more weight to *RE* and reduces to zero the values considered with more collinearity or the covariates with higher VIF values. In Ridge, the coefficient value of *RE* is also larger than the rest of covariates, this means Ridge assigns more weight to *RE* relative to the other independent variables. Instead of assigning zero values to covariates with less weight, in Ridge regression smaller values are assigned to the other covariates. These results confirm that Ridge is in some sense in agreement with LASSO regression coefficients.

Even though the study has a small number of covariates, the multivariable statistics approach emphasizes that renewable energy covariate possesses more weight than other covariates. According to this result the more renewable energy generation a country has, the more electricity consumption per capita exists (Figure 2). This statement poses the question of whether (i) a country with a high percentage of renewable energy generation tends to consume more electricity because their environmental impact is less than a fossil fuel-based model, (ii) since the country has a high electricity consumption, they need to invest more in renewable energy generation to fulfill the electricity demand requirements, or (iii) countries that are in rapid growth of *EC* have more *RE* because they were developing in an era that *RE* was more affordable, or of higher climate awareness.

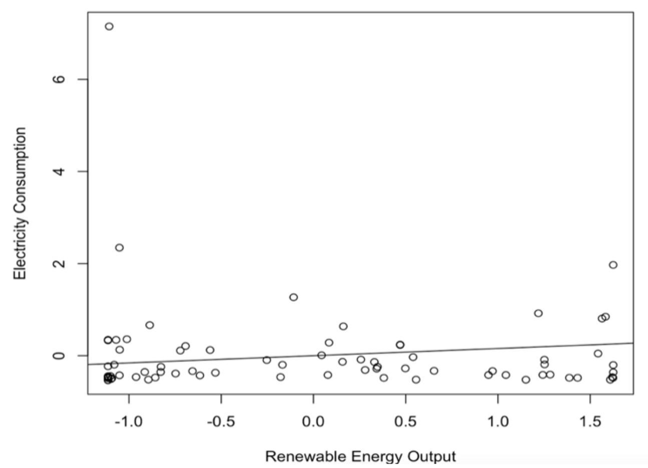

Figure 2. Renewable Energy Output vs. Electricity Consumption

This data demonstrates that energy policy makers at the national and international level need to consider that solving the problem of energy access requires a holistic examination of multiple factors (e.g., social, economic, environmental, technical).

5. DISCUSSION

The multivariable statistical approach to energy access confirms how a penalized regression, such as LASSO and Ridge, assigns more weight to the renewable energy generation covariate in terms of electricity consumption per capita in countries with medium and low HDI. Contrary to OLS, LASSO and Ridge regressions reduced to zero the coefficients attempting to avoid values with high collinearity.

According to this model, renewable energy has the potential to contribute to the increase of electricity consumption in the countries of the Global South, and thus increase energy access rates. Responsible renewable energy initiatives can be instrumental to promote new job opportunities, combat climate change, and increase energy access rates [29]. The increase of energy access rates helps to reduce gender inequality, dynamize rural local economies, and achieve better education and health systems [30]. As this study shows, investing in renewable energy projects could help in reducing energy poverty rates in the countries of the Global South.

Having a robust renewable energy mix that harnesses local sources is beneficial for the countries of the Global South, since this energy mix could help these countries to achieve a certain degree of energy autonomy. Moreover, having energy autonomy would enable countries to invest more in electrical infrastructure thus achieving a higher energy access rate. By transitioning to renewable energy, the countries of the Global South that rely on fossil fuels to generate electricity could reduce their dependency on international oil markets. Conventional renewable energy technologies are more affordable than a decade ago. Despite this economic opportunity, the countries of the Global South have difficulties making an energy transition towards these clean technologies.

For the extended version of this study, it is recommended to examine more socio-economic and environmental variables. A dataset with more covariates would strengthen the LASSO and Ridge covariate weights.

6. CONCLUSSION

From an energy perspective, this study is relevant because it demonstrates the correlation between electricity consumption and renewable energy output in a Global South scenario. The model shows that the higher the presence of renewable energy generation, the higher the electricity consumption.

Achieving greater energy access rates in rural communities of the Global South requires a multi-dimensional approach. Using an OLS helps to examine different variables, and determine which values are more correlated to electricity consumption per capita. To reduce complexity and prevent over-fitting from OLS, this study employed LASSO and Ridge regressions obtaining that *RE* has more weight than the other independent variables.

The use of multi-dimensional variables such as socioeconomic, environmental, technical, enriched the scope of the analysis. A macro level assessment of socioeconomics, environmental and energy variables helped to understand the behavior of electricity consumption within a more comprehensive context. This study could be a tool for energy planners in grasping the importance of renewable energy installations in the countries of the Global South.

The installation of more renewable energy projects around the world, especially with a rural electrification emphasis, is important because such projects improve the quality of life for the world's most vulnerable people. Renewable energy is pivotal to reduce the effects of global warming and to combat climate change. The promotion of renewable energy with an energy justice emphasis is vital for a post-pandemic recovery.

REFERENCES

[1] International Energy Agency (IEA). (2019). SDG 7: Data and Projections. Retrieved January 9, 2021, from www.iea.org/reports/sdg7-data-and-projections

[2] International Renewable Energy Agency (IRENA). (2021). Universal Access to Sustainable Energy Will Remain Elusive Without Addressing Inequalities. Retrieved July 10, 2021, from www.irena.org/newsroom/pressreleases/2021/Jun/Universal-Access-to-Sustainable-Energy-Will-Remain-Elusive-Without-Addressing-Inequalities

[3] Riva, F., & Colombo, E. (2020). System-dynamics modelling of the electricity-development nexus in rural electrification based on a Tanzanian case study. *Energy for Sustainable Development*, 56, 128-143.

[4] Van Ruijven, B. J., Schers, J., & van Vuuren, D. P. (2012). Model-based scenarios for rural electrification in developing countries. *Energy*, 38(1), 386-397.

[5] Siksnelyte-Butkiene, I., Streimikiene, D., Lekavicius, V., & Balezentis, T. (2021). Energy poverty indicators: A systematic literature review and comprehensive analysis of integrity. *Sustainable Cities and Society*, 102756.

[6] Kanagawa, M., & Nakata, T. (2008). Assessment of access to electricity and the socio-economic impacts in rural areas of developing countries. *Energy policy*, 36(6), 2016-2029.

[7] Wassie, Y. T., & Adaramola, M. S. (2021). Socio-economic and environmental impacts of rural electrification with Solar Photovoltaic systems: Evidence from southern Ethiopia. *Energy for Sustainable Development*, 60, 52-66.

[8] Aglina, M. K., Agbejule, A., & Nyamuame, G. Y. (2016). Policy framework on energy access and key development indicators: ECOWAS interventions and the case of Ghana. *Energy Policy*, 97, 332-342.

[9] Ministerio de Energía y Minas (MEM). (2019). Política Nacional de Electrificación Rural 2019-2032. Retrieved May 17, 2021, from www.mem.gob.gt/wp-content/uploads/2018/11/Pol%C3%ADtica-Electrificación-Rural-2019-2032.pdf
[10] Empresa Nacional de Transmisión Eléctrica (ENATREL). (2019). Cobertura Eléctrica del Departamento de Jinotega. Retrieved May 17, 2021, from www.enatrel.gob.ni/jinotega
[11] Instituto Nacional de Estadística e Informática (INEI). (2017). Perú: Perfil Sociodemográfico. Informe Nacional. Retrieved May 17, 2021, from www.inei.gob.pe/media/MenuRecursivo/publicaciones_digitales/Est/Lib1539/
[12] International Renewable Energy Agency (IRENA). (2012). Renewable Energy Technologies: Cost Analysis. Retrieved August 20, 2021, from www.irena.org/documentdownloads/publications/re_technologies_cost_analysis-hydropower.pdf
[13] Costantini, V., & Monni, S. (2008). Environment, human development and economic growth. Ecological Economics, 64(4), 867-880.
[14] United Nations Development Program (UNDP). (n.d.). Human Development Index (HDI). Retrieved November 1, 2021, from hdr.undp.org/en/content/human-development-index-hdi
[15] Organization for Economic Co-operation and Development (OCDE). (n.d.). Linking Renewable Energy to Rural Development. Retrieved July 21, 2021, from www.oecd.org/regional/regional-policy/Renewable-rural-energy-summary.pdf
[16] Kaygusuz, K. (2012). Energy for sustainable development: A case of developing countries. Renewable and Sustainable Energy Reviews, 16(2), 1116-1126.
[17] Intergovernmental Panel on Climate Change (IPCC). (2020). Energy is at the heart of the solution to the climate challenge. Retrieved October 20, 2021, from www.ipcc.ch/2020/07/31/energy-climatechallenge/
[18] The World Bank (WB). (2020). Population, total. Retrieved September 29, 2021, from data.worldbank.org/indicator/SP.POP.TOTL?view=chart
[19] Energy Information Agency (EIA). (2018). Electricity- Electricity Net Consumption (billion kWh). Retrieved October 1, 2021, from www.eia.gov/international/data/world/electricity/electricity-consumption
[20] The World Bank (WB). (2015). Renewable energy consumption (% of total final energy consumption). Retrieved March 13, 2021, from data.worldbank.org/indicator/EG.FEC.RNEW.ZS?view=chart
[21] The World Bank (WB). (2018). Population growth (annual %). Retrieved March 13, 2021, from data.worldbank.org/indicator/SP.POP.GROW?view=chart
[22] The World Bank (WB). (2018). GDP per capita (current US$). Retrieved March 13, 2021, from data.worldbank.org/indicator/NY.GDP.PCAP.CD?view=chart
[23] The World Bank (WB). (2020). Rural population (% total of population). Retrieved August 11, 2021, from data.worldbank.org/indicator/SP.RUR.TOTL?view=chart
[24] The World Bank (WB). (2020). School enrollment, tertiary (% gross). Retrieved March 13, 2021, from data.worldbank.org/indicator/SE.TER.ENRR?view=chart
[25] The World Ban (WB). (2018). CO2 emissions (metric tons per capita). Retrieved September 25, 2021, from data.worldbank.org/indicator/EN.ATM.CO2E.PC?view=chart
[26] Abdi, H., & Williams, L. J. (2010). Normalizing data. *Encyclopedia of research design*, 1.
[27] Daoud, J. I. (2017, December). Multicollinearity and regression analysis. In *Journal of Physics: Conference Series* (Vol. 949, No. 1, p. 012009). IOP Publishing.
[28] Owen, A. B. (2007). A robust hybrid of lasso and ridge regression. *Contemporary Mathematics, 443(7),* 59-72.
[29] Gielen, D., Boshell, F., Saygin, D., Bazilian, M. D., Wagner, N., & Gorini, R. (2019). The role of renewable energy in the global energy transformation. *Energy Strategy Reviews*, 24, 38-50.
[30] Alstone, P., Gershenson, D., & Kammen, D. M. (2015). Decentralized energy systems for clean electricity access. *Nature Climate Change*, 5(4), 305-314.

